

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2005, 11:22:19 ; Search time 163 Seconds (without alignments)

1 MDPTTLALSLLAACVLAGCS.....YKFYDNKRIDSSTGHHHHHH 464

Title: US-09-719-379A-81

Perfect score: 2477

Sequence: Scoring table: BLOSUM62 Gapop 10.0 , Gapext 0.5

Scanned: 2105632 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database : A_Geneseq_16D004:*

- 1: geneseqD19808:*
- 2: geneseqD19908:*
- 3: geneseqD20008:*
- 4: geneseqD20018:*
- 5: geneseqD20028:*
- 6: geneseqD20038:*
- 7: geneseqD20038:*
- 8: geneseqD20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No. Score Query Match Length DB ID

1 2477 100.0 464 3 AAY79993

2 1920 77.5 364 2 AAB20108

3 1251.5 50.5 358 6 ABB84614

4 1219.5 50.5 359 7 ABO061848

5 1210.5 49.2 364 7 ADF05718

6 1210.5 48.9 359 6 ABR68758

7 1074 43.4 356 2 AAY52773

8 661.5 26.7 451 6 ABU04447

9 661.5 26.7 451 6 ABU04447

10 651 26.3 144 2 AAY02348

11 606 24.5 446 2 AAY06590

12 606 24.5 324 2 AAY02352

13 606 24.5 324 4 AAG63234

14 606 24.5 324 7 ADB67693

15 597 24.1 326 2 AAY02354

16 597 24.1 326 4 AAG63236

17 597 24.1 326 7 ADB67697

18 594.5 24.0 411 4 AAG63235

19 594.5 24.0 411 7 ADB67695

20 588.5 23.8 411 2 AAY02353

21 585.5 23.6 413 2 AAY02355

22 585.5 23.6 413 4 AAG63237

23 585.5 23.6 413 7 ADB67699

24 565 22.8 220 2 AAY25375

25 565 22.8 220 2 AAY02631

26 565 22.8 220 2 AAY02634

27 565 22.8 421 5 AAM50663

28 565 22.8 421 5 AAY25385

29 563 22.7 278 2 AAY02641

30 563 22.7 278 2 AAY25384

31 561.5 22.7 227 2 AAY02640

32 561.5 22.7 227 2 AAY25376

33 559.5 22.6 273 2 AAY25376

34 559.5 22.6 273 2 AAY02632

35 559.5 22.6 371 2 AAY25377

36 559.5 22.6 371 2 AAY02633

37 558 22.5 220 8 ADL64005

38 558 22.5 220 8 ADL63966

39 558 22.5 220 8 ADL97335

40 558 22.5 227 2 AAY25382

41 558 22.5 227 2 AAY02638

42 558 22.5 383 2 AAY25386

43 558 22.5 383 2 AAY02642

44 429.5 17.3 797 4 ABG29333

45 381.5 15.4 386 6 ADA34227

ALIGNMENTS

RESULT 1
AAY79993 standard; protein; 464 AA.

ID XX
AC XX
DT 15-MAY-2000 (first entry)
XX
DB Plasmid LPD-LB1-III protein sequence.

XX
Vaccine; non-typeable *Haemophilus influenzae*; nHi; infection; chimeric protein; *Haemophilus influenzae*; P5-like fimbrial protein; chimeric protein D; LBL(f); immunogen; antigenic; otitis media; sinusitis; conjunctivitis; lower respiratory tract infection.

XX
Haemophilus influenzae.

OS

Synthetic.

PN WO9964067-A2.

XX

XX
PD 16-DEC-1999.
XX
PF 28-MAY-1999;
XX
PR 11-JUN-1998;
XX
PA (SMIK) SMITHLINE BECHMAN BIOLOGICALS.
PA (OHIS) UNIV OHIO STATE RES FOUND.
XX
PI Bakalatz LO, Cohen J, Dequesne G, Labet Y;
XX
DR WPI ; 2000-116457/10.
DR N-PSDB; AAZ91252.

XX
PT Novel antigenic P5-like fimbrial subunit peptides used in vaccines against *Haemophilus influenzae*.
XX
PS Claim 14: Fig 5; 60pp; English.
XX
CC The present invention describes antigenic P5-like fimbrial subunit peptides (LBL(f) peptides) of P5-like fimbrial proteins from various *Haemophilus influenzae* strains. The peptides are used for diagnosis, prevention, and treatment of *Haemophilus influenzae* infections, such as otitis media, sinusitis, conjunctivitis, or lower respiratory tract infection. The peptides may also be used in vaccines against H. influenzae. Antibodies and probes from the present invention can be used for diagnosis of H. influenzae infection. AAY79993 to AAY79993, and AAZ91201 to AAZ91252, represent sequences used in the exemplification of

acquired immunodeficiency syndrome; neurodegenerative disease; stroke; Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis; spinal muscular atrophy; retinitis pigmentosa; cerebellar degeneration; myelodysplastic syndrome; ischaemic injury; myocardial infarction; myocardial infarction; liver disease; idiopathic dilated cardiomyopathy; ischaemic cardiomyopathy; aplastic anaemia; chronic neutropenia; mania; senile dementia; Huntington's disease; hypertension; anxiety; severe bipolar affective disorder; glycerocephosphoryl phosphodiester phosphodiesterase.

XX OS Escherichia coli.

PN WO200281516-A2.

XX PD 17-OCT-2002.

XX PP 16-JAN-2002; 2002WO-US001098.

XX PR 16-JAN-2001; 2001US-0262306P.

PR 15-JAN-2002; 2002US-00047855.

XX PA (MILLI-) MILLENNIUM PHARM INC.

PI Chiang LW;

XX WPI: 2003-058503/05.

XX PT Novel isolated programmed cell death-related polypeptide, NARC10 and NARC16, useful for treating disorders associated with abnormal apoptotic process, e.g. Alzheimer's disease, cancer, myocardial infarction, stroke.

XX PS Disclosure; Fig 3; 123pp; English.

XX This invention describes novel cell death-related polypeptides NARC10 and NARC16, located on chromosome 4q11-4q21 and which have cardiotropic, antiHIV, immunosuppressive, dermatological, antiinflammatory, cerebroprotective, neuroprotective, antianemic, cardiotropic, vasotropic, antianamic, antidiabetic, immunosuppressive, cyclostatic, thymomimetic, nephrotropic, immunosuppressant, tranquiliser, hypotensive and neuroleptic activity and can be used in gene therapy. The products of the invention can be used to modulate NARC10 or NARC16 polypeptides or polynucleotides, to map NARC genes on a chromosome, e.g. to locate gene regions associated with genetic disease or to associate NARC10 or NARC16 with a disease. The polypeptides are also useful for modulating the apoptotic process, and are therefore useful for modulating, and treating disorders associated with increased apoptosis, inhibition of apoptosis or disruptions in cell cycle, for regulating cellular functions including programmed cell death, nucleosome assembly, phosphate homeostasis and the cell cycle. Preferably, the products of the invention are useful for treating disorders associated with abnormally low rate or abnormally high rate of apoptosis, e.g. cancers including follicular lymphomas, carcinomas with p53 mutations, or hormone-dependent tumours, autoimmune disorders including systemic lupus erythematosus, diabetes, graft rejection, Hashimoto's thyroiditis and immune-mediated glomerulonephritis and viral infections e.g. infections caused by herpes viruses, virus-induced lymphocyte depletion (including acquired immunodeficiency syndrome (AIDS)), neurodegenerative diseases manifested by loss of specific sets of neurons (including Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis, spinal muscular atrophy, retinitis pigmentosa, and cerebellar degeneration), myelodysplastic syndromes (including aplastic anaemia), ischaemic injuries (including myocardial infarction, stroke and reperfusion injury), and toxin (e.g. alcohol) induced liver disease, idiopathic dilated cardiomyopathy, ischaemic cardiomyopathy and valvular heart disease, aplastic anaemia, chronic neutropenia, and myelodysplastic syndromes, central nervous system disorders, senile dementia, Huntington's disease, hypertension, schizophrenia, attention deficit disorder, mania, anxiety, severe bipolar affective disorder (BP-1). This sequence represents the E. coli glycerocephosphoryl phosphodiester phosphodiesterase protein described in the method of the invention.

XX Sequence 358 AA;

Query Match 50.5%; Score 1251.5; DB 6; Length 358;
Best Local Similarity 64.1%; Pred. No. 2.7e-100;
Matches 56; N1 matches 65; Indels 7; Gaps 2;

Matches 229; Conservative 4 KTLASLLANGVLAGSSHSSNNMANTQMSDKIIIAHGRASGYLDEHTLSEKALPAQQA 63
6 KNLSMAMMSTVMSANAAAD----SNEKIVIAHGRASGYLPEHTLPAKAMAYAQQA 59

Qy 64 DYLEQDLMPTKDGRIVVTHDFLDLDTYAKKPPHRKDGRYYTDFTRKEIQSDEMTE 123
Db 60 DYLQDLMVTKDDNLYLVDHYLDVTDYADREPRAKRDGRYTAIDPFLDEIKSFKTE 119

Qy 124 NFETIKDGKQAOVYPKRFPLWKSHPFRIHTPFDDEBIEPQGLEKSTGKVGTYPEIKAPWFHH 183
Db 120 GFDINGKVKVQTPRFPMSDFPVHTPFBIEIVQGLNHSITGNIGTYPEIKAPWFHH 179

Qy 184 QNGKQIAAETLKULKKYGYDKKTDNYLQTPDFNELKRIKTELLQMGMDKLVQLIAYT 243
Db 180 QBGKQDAAKTLLEVTKKYGTVGKQDFADLKLKNEELPDRGMELNLVQLIAYT 239

Qy 244 DWKETQEKDQDKGKQYWNKQYNTDMFKRQAMAEVKYTAQGVGQWMMVUNKESKPKDNPVYTP 303
Db 240 DWNETQKQDPDGSWNTYNDMMFKRQAMQVAEFDGQDGYHMLI-BETSQGKNIKLTG 298

Qy 304 LVKELLAQYNVEVHPPYTVRKDALPEFTDYNQMDALLNSGATGVFTDFTGVRFL 360
Db 299 MVQDAQQNKLVWHPVTVRSQDKLPEVTPDVNQLYDALYNKAGVNGIQLTDFDKAVRFL 355

RESULT 4
ID ABO61848 standard; protein; 359 AA.

XX ABO61848; AC ABO61848; DT 29-JUL-2004 (first entry)

XX Klebsiella pneumoniae polypeptide seqid 8365..

XX Recombinant expression vector; transcription regulatory element; Klebsiella pneumoniae protein; antibacterial; Vaccine.

XX OS Klebsiella pneumoniae.

XX PN US6610836-B1.

XX PD 26-AUG-2003.

XX PP 27-JAN-2000; 2000US-00489039.

XX PR 29-JAN-1999; 99US-011747P.

XX PA (GENO-) GENOME THERAPEUTICS CORP.

XX PI Bretton GL, Osborne M;

XX DR WPI: 2003-89546/82.

XX DR N-PSB; ACH95399.

XX PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for preparing a vaccine composition against Klebsiella pneumoniae.

XX PS Disclosure; SEQ ID NO 8365; 932pp; English.

XX The invention describes new isolated nucleic acid encoding a Klebsiella pneumoniae polypeptide. Also described are: a recombinant expression vector comprising the nucleic acid, operably linked to a transcription regulatory element, and a cell comprising the recombinant expression vector. The nucleic acid is useful for preparing a vaccine composition against Klebsiella pneumoniae. This is the amino acid sequence of a Klebsiella pneumoniae polypeptide of the invention

XX

SQ	Sequence 359 AA;	CC polypeptide and a method for screening test compounds for anti-bacterial activity. The polypeptides and polynucleotides are useful as molecular targets for diagnosing, preventing and treating pathological conditions resulting from bacterial infection, as reagents for diagnosis of bacterial diseases, as components of antibacterial vaccines, or as targets for antibacterial drugs or as bio-control agents for plants. This sequence represents a <i>Proteus mirabilis</i> polypeptide of the invention.
Query Match	50.5%; Score 1251.5; DB 7; Length 359;	CC
Best Local Similarity	65.3%; Pred. No. 2.7e-100;	CC
Matches 231; Conservative	48; Mismatches 74; Indels 1; Gaps 1;	CC
Qy	7 ALSLLAAGVLLAGCSSHSSNMANANTQMSKDSKIIIAHRAASGYLPEFTLESKALAFAAQADYL 66	CC
: Db	8 AMKMTKALMSGMISSALACSATADKVIYAHRAASGYLPEFTLESKALAFAAQADYL 62	CC
SQ	Sequence 364 AA;	CC
Query Match	49.2%; Score 1219.5; DB 7; Length 364;	CC
Best Local Similarity	64.7%; Pred. No. 1.7e-97;	CC
Matches 229; Conservative	50; Mismatches 70; Indels 5; Gaps 3;	CC
Qy	8 ISLJAAAGTLAGCSCYHSNNMANANTQMSKDSKIIIAHRAASGYLPEFTLESKALAFAAQADYL 67	CC
: Db	12 LKPLVAGVLTLS-LSTIA-QASDKVVIYAHRAASGYLPEFTLESKALAFAAQADYL 67	CC
Qy	127 TDKGKQAOQYENPNBPLWKSHEPRHFTEDDEIEFQLEKSTGKVGTYPEIKAPNPHQNG 186	CC
: Db	123 PRNGKNVQTYTPGRFPNCKSDPRHFTEEIFVQGLNHSCKTKNIGTYPEIKAPNPHQNG 182	CC
Qy	187 KDIAAETLKLVKKKYGDKTKDMVYLQTFDFNBLKRKTELLPQMGMNDLKVLYQLIAYTDW 246	CC
: Db	183 KDIAAETLKLVKKKYGDKTKDMVYLQTFDFNBLKRKTELLPQMGMNDLKVLYQLIAYTDW 242	CC
Qy	247 ETQEKDPKGKWWNNYNTDMWFKPGAMAEVVKYADGVGPWYMLVNLKGEESKEDPNVITYTPLYK 306	CC
: Db	243 ETQKQADGKWWNSDWMFKPGAMAQIAQADGKDYHMLV-AEGSKPGCAVLTAMW 301	CC
Qy	307 ELAQXNVBPHPTVTKDALLPEFTDYNQMDALLNSGATGATVFTDPTGVEFL 360	CC
: Db	302 EAHASHLQVHPPTVRAQDPLQPEVATNQNQLYDVLQAGVDSLFTFPDKAVQFL 355	CC
RESULT 5	248 TQEKEPKGKWWNNYNTDMWFKPGAMAEVVKYADGVGPWYMLVNLKGEESKEDPNVITYPLVKE 307	CC
ID ADP05718	248 TYRSPDPGTWNTNSYDMMWFKPGAKMKEIATADGIPDQYHMLV-EBDSTDPKITLGMAAD 306	CC
XX	248 ABM68758 standard; protein; 359 AA.	CC
AC ADP05718;	248 ABM68758;	CC
XX	20-NOV-2003 (first entry)	CC
DT 12-FEB-2004 (first entry)	20-NOV-2003 (first entry)	CC
DB	248 ABM68758	CC
XX	Photorhabdus luminescens protein sequence #1855.	CC
KW	Antibacterial; fungicide; insecticide; polymorphism; genetic analysis; detection; food; gene expression; plant; animal; microorganism; toxin; antibiotic; biopesticide; virulence factor; disease model; plague; whooping cough.	CC
XX	308 LAQTNVEVPHPTVTKDALLPEFTDYNQMDALLNSGATGATVFTDPTGVEFL 361	CC
AC	307 ARANKLTIDHPFTVRLPKYADGQDLYDIYNOQAGBEGVFTDPDLGVKFLQ 360	CC
XX	307 ARANKLTIDHPFTVRLPKYADGQDLYDIYNOQAGBEGVFTDPDLGVKFLQ 360	CC
OS Photorhabdus luminescens.	OS Photorhabdus luminescens.	CC
XX	OS WO200294867-A2.	CC
PI	XX	CC
XX	WPI; 2003-89291/82.	CC
DR	XX	CC
N-PSDB; AD051546	XX	CC
XX	28-NOV-2002.	CC
PD 12-AUG-2003.	XX	CC
XX	07-FEB-2002; 2002WO-IB003040.	CC
PR 05-APR-2000; 2000US-00543681.	XX	CC
XX	07-FEB-2001; 2001FR-00001659.	CC
PR 09-APR-1999; 99US-0128706P.	XX	CC
XX	(INSP) INST PASTEUR.	CC
PA (GENO-) GENOME THERAPEUTICS CORP.	PA (CNRS) CNRS CENT NAT RECH SCI.	CC
XX	XX	CC
PI Breton GL;	XX	CC
XX	XX	CC
WPI; 2003-89291/82.	XX	CC
DR	XX	CC
N-PSDB; AD051546	XX	CC
XX	XX	CC
New <i>Proteus mirabilis</i> polypeptides and polynucleotides, useful as reagents for diagnosis of bacterial disease, as components of antibacterial vaccines, as targets for antibacterial drugs, or as biocontrol agents for plants.	XX	CC
Disclosure; SEQ ID NO 6003; 870pp; English.	XX	CC
The invention relates to new <i>Proteus mirabilis</i> polypeptides and polynucleotides. The invention also relates to antibodies against the polypeptides. The invention also relates to methods of producing, methods for immunising and individual against <i>P. mirabilis</i> , a method for generating vaccines for immunising an individual against <i>P. mirabilis</i> , a method for evaluating a compound for the ability to bind a <i>P. mirabilis</i>	CC	

useful e.g. as therapeutic antimicrobials and agricultural pesticides.

PTX XX Claim 2; SEQ ID NO 1855; 1205pp; Franch.

CC The invention relates to the isolation of genes and their encoded proteins from *Photobacterium* *luminosens*. The isolated sequences are sources of probes and primers for detecting the genome of *P. luminosens* and related species; to study polymorphisms; for gene analysis and for detection/amplification of the genes. Antibodies (Ab) raised against the polypeptides encoded by the genes are used for detection/identification of *P. luminosens*, e.g. in foods. The genes (proteins, Ab and cells that carry a gene-containing vector are used to select compounds that modulate, regulate, induce or inhibit expression of the genes in plants, animals or microorganisms other than *P. luminosens* and are able to alter response or sensitivity to toxins and antibiotics produced by *P. luminosens*. Cells transformed to express the genes are useful for recombinant production of the proteins, particularly toxins and antibiotics useful as insecticides, bactericides and fungicides. The genes, proteins, vectors containing the genes and Ab are also useful therapeutically (to treat microbial infection by bacteria or fungi that are sensitive to *P. luminosens*-encoded toxins or antibiotics) and as biopesticides. Other uses of the genes and the proteins are as virulence factors and for identifying targets of human diseases for which *P. luminosens* is a model (particularly plague and whooping cough). This sequence represents one of the isolated *P. luminosens* proteins

Sequence 359 AA;

Query Match 48.9%; Score 1210.5; DB 6; Length 359;
Best Local Similarity 62.0%; Pred. No. 1e-96;
Matches 222; Conservative 54; Mismatches 71; Gaps 2;

Qy 4 KTLALSLLAGVLAGCSSSHSSMANTOMKSIIIAHGRGASCYLPBHTLESKALAPQA 63
Db 6 KTMIGIL-----TSSMSGTAQAAKIVIAHGRGASCYLPBHTLPKAMAYEGA 55

Qy 64 DYLEQDLAMTIDGRLVYTHDHFGLDTWACKCPHRHRKDRGRRYVIDFTLKEIOSLME 123
Db 56 DYLEQDLWTKDDELVHDYDRLDVTWAKCPHRHRKDRGRRYVADFTLSEISLKPTE 115

Qy 124 NFETKDQKQAOQVYQPNRFPFLWKSHERIHTFEDRIEPIQGLEKSTGKVGTYPEPIKAPWFHH 183
Db 116 GFDIKNDRQIQNFSNRFPFLWKSDFRHTFQERIEFWGLNSTGKNGTYPEPIKAPWFHQ 175

Qy 184 QNGKDIATELTKVLKKKGYDKCTDMVLYLQTDPFENELKRKTRTELLPQMDKLKVQLIAYT 243
Db 176 KEGKDIDSKVLAQVKGYTKSDFKVIQOCPTNELKRNKELEPKLGMDKLKVQLIAYT 235

Qy 244 DWEKTOERKDPKGKSYWNTYNDMFKPGKAMAEVYKYADCGPGWYMLNEKEESKRPDNIVYTP 303
Db 236 DNNETYEWQSDGKWTWNSYDWNFKPGKAMKEVAQADGDPYTHMLVEK-YSTPTNKLIN 294

Qy 304 LYKELAQYNEVHPYTYRVDALPEFTTDVNOMYDALINKSGATGVFTDFPDTGVEFLK 361
Db 295 LYKEAHTNQNLEVHPYTYRVDQLPKYATSGDQLFDIYNAQGYDGVFTDFPDLGKFLQ 352

/

RESULT 7

AAY52773 standard; protein; 356 AA.
ID AAY52773 .
XX AAY52773; .
AC .
XX .
XX 26-JAN-2000 (first entry)

DT .
XX .
DB .
Treponema pallidum glycerophosphodiester phosphodiesterase.

XX Treponema pallidum; syphilis; vaccine; treponemal disease; Msp; Yaws; major sheath protein; bejai; gingivitis; periodontal disease; pinta.
XX Treponema pallidum.
XX WO9953099-A1.

RESULT 8
AAY06589

XX 21-OCT-1999.
PS XX .
XX 09-APR-1999; 99WO-US007886.

CC XX .
CC 10-APR-1998; 98US-00058968.

CC XX (UNIV) UNIV WASHINGTON.

CC XX Van Voorhis WC, Lukehart SA, Centurion-Lara GA, Cameron CES;

CC XX WPI; 1999-62045/53.
N-PSDB; AAZ33104.

XX XX Novel proteins useful in vaccines against syphilis and other treponemal diseases.

XX XX PS Claim 1; Page 94-95; 200pp; English.

CC The present invention describes novel *Treponema pallidum* genes. These genes encode a glycerophosphodiester phosphodiesterase (Gpd), a D15/Cma87 homologue, and proteins with homology to major outer sheath (Msp) proteins of *T. denticola*. Also described are: (1) an isolated protein capable of inducing a protective immunologic response to *T. pallidum*, *T. pertenue*, or *T. pallidum*, when administered in an effective amount to an animal host; (2) a method of identifying a *T. pallidum* vaccine candidate; (3) a method of inducing a protective immune response against *T. pallidum*; (4) a method for obtaining a sample of DNA to determine whether it originated from *T. p.* subspecies *pallidum*, *T. p.* subspecies *pertenue* or *T. p.* subspecies *endemicum*, and (5) a method of determining whether a first and a second clinical isolate of *T. p.* *pallidum* are the same or different. The proteins are used, either alone or in combination, in vaccines against Treponemal diseases, e.g. syphilis, bejai, pinta, yaws, gingivitis, and periodontal disease. They may also provide protection against other Treponemal diseases. The methods may be used to identify vaccine candidates, and to determine the origin of a treponemal nucleic acid. AAY52773 to AAZ33104 and AAY52773 to AAY52831, represent sequences used in the exemplification of the present invention.

XX Sequence 356 AA;

SQ .

Query Match 43.4%; Score 1074; DB 2; Length 356;
Best Local Similarity 54.6%; Pred. No. 8.1e-85;
Matches 190; Conservative 67; Mismatches 79; Indels 12; Gaps 1;

Qy 13 AGVLAGCSSHSSNMANQMSDKIITAHGRGASGYLPEHTLESKALAFQAOQADYLEQDLM 72
Db 15 AALVAGCA-----SERMVAYRGAGTYPEHTPAVKVLAQGADYLOQDGVVL 62

Qy 73 TKDGHLVVHDHFGLQCLTVAKKPFPHRHRKDRGYYVUDFLIKEIQLSLENTEFETKDGKQ 132
Db 63 SKDNQOLIVAQSHILNMDTDAEKPFRRQRDAGHPVWIDFTVEELSLRATNSFYTRGKRH 122

Qy 133 AQVYVNPFRFLWKSFRHTFDEIETQGLKSTKRGVGYPEIKAQWPHFHQNGDIAAE 192
Db 123 TPVIGORFLPKWPKGFLRHTFEEQFLIGLEOTGKIGIYSEIKVUPWFFHQEGDIAAL 182

Qy 193 TLKVLLKKYGYDKKTDQVYDFDNEKLKRKTRTELLPQMGMDKLKVQLIAYT 252
Db 183 TLALIJKYGYQSRSLVYVOTYDFNBLKRTRKELLPKYNKVLQVRAVTDQRSRQEQD 242

Qy 243 SRGKVNNTNNWMMFPGKQKIAKJADGVGPMVSLQDQK 302
Db 313 VEHPPVTVRDLAPPEFTDYNQMYDALINKSGATGVFTDFPDTGFBFL 360
Qy 303 LECHHTVTRKELLPSPYARTMDEMFSTLKFQGTVANVVLTDPEFLGKFL 350

QY	AAV06589	standard; protein;	451	AA.
AAV06589;				
XX				
XX	26-OCT-1999	(first entry)		
XX	DB	Lipoprotein D-MAGE-3-His fusion protein.		
XX	XX	MAGE-3; lipoprotein D; LPD-MAGE-3-His; fusion protein; tumour; melanoma; breast cancer; bladder cancer; lung cancer; head and squamous cell carcinoma; colon cancer; oesophagus carcinoma; vaccine; human.		
XX	XX	Haemophilus influenzae.		
XX	XX	Homo sapiens.		
XX	XX	Synthetic.		
XX	XX	Chimeric.		
Key		Location/Qualifiers		
Misc-difference	443	/note= "this residue is additional to the residues deduced from the nucleotide sequence of AXR87588"		
FT				
FT				
FT				
XX	XX	W09340186-A2.		
PR	PN	XX		
PR	PD	12-AUG-1999.		
PR	PF	02-FEB-1999;	99WO-EP000660.	
PR	PF	05-FEB-1998;	98GB-00002543.	
PR	PR	06-FEB-1998;	98GB-00002650.	
XX	XX	(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.		
PI	PI	Cabazon Silva T, Cohen J, Slaoui MM, Vinals Bassols C;		
XX	XX	WPI: 1999-494293/41.		
XX	XX	N-PSDB; AAX87588.		
DR	DR			
DR	DR			
XX	XX	New protein derivatives used in cancer vaccine therapy for treating a range of cancers including melanomas, carcinomas and cancers of breast.		
XX	XX	Example 1; Page 64-65: 72pp; English.		
CC	CC	This sequence represents a novel fusion protein composed of lipidated protein D (LpD) of Haemophilus influenzae B, the human MAGE-3 tumour-associated antigen and a hexahistidine tail. A vector designed for recombinant expression of the fusion protein is provided. MAGE-3 cDNA was amplified using primers that altered the first 5 codons to Escherichia coli codon usage. The LpD moiety provided the fusion protein with additional exogenous T-cell epitopes and also increased expression levels in E. coli. The lipid tail ensured optimal presentation of the antigen to antigen-presenting cells. The affinity tag facilitated purification. The invention relates to MAGE proteins fused to an immunological fusion partner, e.g. LPD-MAGE-3-His. These novel fusion proteins provide vaccines for immunotherapy of melanomas or other MAGE-associated tumours like breast, bladder, lung and non-small cell lung cancer, head and squamous cell carcinoma, colon carcinoma and oesophagus carcinoma		
CC	CC	Sequence 451 AA;		
CC	CC	Query Match 26.7%; Score 661.5; DB 2; Length 451;		
CC	CC	Best Local Similarity 37.5%; Pred. No. 1..1e-48;		
CC	CC	Matches 194; Conservative 51; Mismatches 152; Indels 121; Gaps 18;		
CC	CC	1 MDPKTLALSLIAAGTLAGCSHSSMANTQMSKDIIIAHRGAGSGLPHTLESKALAF 60		
CC	CC	1 MDPKTLALSLIAAGTLAGCSHSSMANTQMSKDIIIAHRGAGSGLPHTLESKALAF 60		
CC	CC	61 QQADYLEQDLMATKDGRLVYTHDPLDGLTIDVAKKFPHHRKGDRYYVDFLKEIOSIE 120		
CC	CC	61 QQADYLEQDLMATKDGRLVYTHDPLDGLTIDVAKKFPHHRKGDRYYVDFLKEIOSIE 120		
SQ	XX	(ZYCO) ZYCO INC.		
XX	XX	Chiciz RM, Tomlinson AJ, Urban RG;		
XX	XX	WPI; 2003-040607/03.		
XX	XX	New polypeptides (e.g. kinases, phosphatases, proteases, transporters, cytoskeletal proteins, receptors or transcription factors), useful for treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or leukemia.		
XX	XX	Example 2; SEQ ID NO 1113; 134pp; English.		

The invention describes a purified polypeptide, which comprises a fragment of a kinase, phosphatase, protease, protease inhibitor, transporter, cytoskeletal protein, receptor or transcription factor. The polypeptide is useful as an immunogenic composition for eliciting in a mammal an immunogenic response directed against any of the purified polypeptide. The purified polypeptide, or the antibody that binds to this polypeptide, is useful for treating cancer. The polypeptide is also useful for identifying compounds that binds to a naturally processed class I MHC binding polypeptide. The polypeptides and nucleotides are particularly useful for treating or preventing myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma, lymphoma or leukemia. These are also useful for screening agents for treating the above mentioned diseases. This sequence represents an expressed protein tag (EPT) isolated from human tissue for translational profiling. Note: This sequence does not appear in the printed specification but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

Sequence 451 AA;

Query Match 26.7%; Score 661.5; DB 6; Length 451;

Best Local Similarity 37.5%; Pred. No. 1.1e-48; Matches 194; Conservative 51; Mismatches 152; Indels 121; Gaps 18;

Qy 1 MDPKTLAISLLAAGVLAGCSSSHSNMANTQMSDKIIIAHRCAGSYLPHEFTLESKALAPA 60

Db 1 MDBKTLAISLLAAGVLAGCSSSHSNMANTQMSDKIIIAHRCAGSYLPHEFTLESKALAPA 60

Qy 61 QQADYLBEDLANTKDGRIVVTHDFLQGLTDVAKKFPRHRKDGRYYVDFTLKEIQSLE 120

Db 61 QQADYLBEDLANTKDGRIVVTHDFLQGLTDVAKKFPRHRKDGRYYVDFTLKEIQSLE 120

Qy 121 MTEENFETDGKQAVQYPNRFLPKMSHFRINTFEDIEBFIQLEKSTGKKGTYPEIKA[PW 180

Db 121 MTEENFETDGKQAVQYPNRFLPKMSHFRINTFEDIEBFIQLEKSTGKKGTYPEIKA[PW 180

Qy 181 FHHQNGKDIQAETTLKVKRKYDVKTDMVYLOTFDFNELKRLTELLPQMGRDL-KLVQL 239

Db 160 TEEQEAASSSSSTLVEV-----TUCEVPAAESPPDPPQSPQGASSLPTT 201

Qy 240 IAYTDWKETOE----KDPKGWVWNVNYDWMFKPG----AMAEVVY-----AD 279

Db 202 MNYPLWCSQSYEDSSNQBEGPSTFDIUSBFOAALSRSKVAELVHFLLLKTRAREPVTKAB 261

Qy 280 GNGW-----GWWMLVNKEEKSXPDKNIVYTPLVKELAQYNVNEVHPYTVRKDALPEFTDVNQMY 336

Db 262 MIGSVVGNWQYFPVTSKASSSLQLVFGIELMEVDPIGHLYI-----PATCIGLSY 313

Qy 337 DALLNKSGATGVTFDPPTGVBFL-KGKSMGGKAGVAFYRDSYK-----TRDHKK 433

Db 314 DGLLGDNQI-----MPKAGLILIVLAIARRGDCAPEKWIWBLSLVEVPEGREDSILG 367

Qy 391 STKLKNGTGHHRARAMDGKAGVAFYRDSYK-----TRDHKK 433

Db 368 DPKCLLTCOH-----FVQENLYLETRQVPSDPACTEFLMGPRALVETSYVKV 413

Qy 434 GRHTARYRSRSDYKFYD-----NKR1DSTSCHHHHHH 464

Db 414 LHHMVKTSGGPHISYPPLHEWVIREGSETSGSHHHHHH 451

RESULT 10

AY02348 standard; protein; 144 AA.

AC AY02348;

XX 17-OCT-2003 (revised)

DT 09-JUL-1999 (first entry)

DE Protein D, the fusion partner for HIV-1 Nef/Tat fusion proteins.

XX HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;

KW vaccine; HIV infection; protein D.

XX Haemophilus influenzae; B.

OS WO916884-A1.

PN 08-APR-1999.

PD 17-SEP-1998; 98WO-EP006040.

PP 26-SEP-1997; 97GB-00020585.

PR (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX PI Bruck C, Godart SAG, Marchand M;

XX DR WPI; 1999-302282/25.

DR N-PSDB; AAX35684.

XX HIV Tat or Nef protein linked to a fusion partner.

XX PT Example 2; Fig 1; 66pp; English.

XX PS The present sequence represents protein D. This protein acts as the fusion partner for the fusion proteins of this invention, which also comprise HIV-1 Nef or Tat proteins (or derivative). The fusion protein can be used in vaccine to prevent HIV infection. (Updated on 17-OCT-2003 to standardise OS field)

CC SQ Sequence 144 AA;

CC Query Match 26.3%; Score 651; DB 2; Length 144;

CC Best Local Similarity 100.0%; Pred. No. 1.7e-48; Mismatches 0; Indels 0; Gaps 0;

CC Matches 127; Conservative 0; Sequence 144 AA;

CC Qy 1 MDPKTLAISLLAAGVLAGCSSSHSNMANTQMSDKIIIAHRCAGSYLPHEFTLESKALAPA 60

Db 1 MDBKTLAISLLAAGVLAGCSSSHSNMANTQMSDKIIIAHRCAGSYLPHEFTLESKALAPA 60

CC Qy 61 QQADYLBEDLANTKDGRIVVTHDFLQGLTDVAKKFPRHRKDGRYYVDFTLKEIQSLE 120

Db 61 QQADYLBEDLANTKDGRIVVTHDFLQGLTDVAKKFPRHRKDGRYYVDFTLKEIQSLE 120

CC Qy 61 QQADYLBEDLANTKDGRIVVTHDFLQGLTDVAKKFPRHRKDGRYYVDFTLKEIQSLE 120

Db 61 QQADYLBEDLANTKDGRIVVTHDFLQGLTDVAKKFPRHRKDGRYYVDFTLKEIQSLE 120

CC Qy 121 MTENFETDGKQAVQYPNRFLPKMSHFRINTFEDIEBFIQLEKSTGKKGTYPEIKA[PW 180

Db 121 MTENFETDGKQAVQYPNRFLPKMSHFRINTFEDIEBFIQLEKSTGKKGTYPEIKA[PW 180

Qy 181 FHHQNGKDIQAETTLKVKRKYDVKTDMVYLOTFDFNELKRLTELLPQMGRDL-KLVQL 239

Db 160 TEEQEAASSSSSTLVEV-----TUCEVPAAESPPDPPQSPQGASSLPTT 201

Qy 240 IAYTDWKETOE----KDPKGWVWNVYDWMFKPG----AMAEVVY-----AD 279

Db 202 MNYPLWCSQSYEDSSNQBEGPSTFDIUSBFOAALSRSKVAELVHFLLLKTRAREPVTKAB 261

Qy 280 GNGW-----GWWMLVNKEEKSXPDKNIVYTPLVKELAQYNVNEVHPYTVRKDALPEFTDVNQMY 336

Db 262 MIGSVVGNWQYFPVTSKASSSLQLVFGIELMEVDPIGHLYI-----PATCIGLSY 313

Qy 337 DALLNKSGATGVTFDPPTGVBFL-KGKSMGGKAGVAFYRDSYK-----TRDHKK 433

Db 314 DGLLGDNQI-----MPKAGLILIVLAIARRGDCAPEKWIWBLSLVEVPEGREDSILG 367

Qy 391 STKLKNGTGHHRARAMDGKAGVAFYRDSYK-----TRDHKK 433

Db 368 DPKCLLTCOH-----FVQENLYLETRQVPSDPACTEFLMGPRALVETSYVKV 413

Qy 434 GRHTARYRSRSDYKFYD-----NKR1DSTSCHHHHHH 464

Db 414 LHHMVKTSGGPHISYPPLHEWVIREGSETSGSHHHHHH 451

RESULT 11

AY02348 standard; protein; 144 AA.

AC AY02348;

XX 17-OCT-2003 (revised)

DT 09-JUL-1999 (first entry)

DE Protein D-MAGE-1-His fusion protein.

XX KW MAGE-1; lipoprotein D; LPD-MAGE-1-His; fusion protein; tumour; melanoma;

OS Homo sapiens.

OS Synthetic.

OS ChimERIC.

XX ID WO940188-A2.

XX DT 26-OCT-1999 (first entry)

DE Lipoprotein D-MAGE-1-His fusion protein.

XX KW MAGE-1; lipoprotein D; LPD-MAGE-1-His; fusion protein; tumour; melanoma;

OS Homo sapiens.

OS Synthetic.

OS ChimERIC.

XX ID WO940188-A2.

XX DT 12-AUG-1999.

DE XX 02-FEB-1999.

XX KW 99WO-EP000660.

Db	301	RHVARELHPEY--FKN---CTSGHHHHHH 324
XX	XX	WO200154719-A2.
XX	XX	02-AUG-2001.
XX	XX	29-JAN-2001; 2001WO-EP000944.
PF	PF	31-JAN-2000; 2000GB-00002200.
PR	PR	14-APR-2000; 2000GB-00009336.
PR	PR	06-JUN-2000; 2000GB-00013806.
PR	PR	28-JUN-2000; 2000WO-EP005998.
XX	XX	HIV-1 Lipop-nef-His amino acid sequence SEQ ID NO:15.
PA	PA	HIV; Tat; HIV-1; Nef; gp120; vaccine; immunisation; human; anti-HIV; HIV fusion.
PI	PI	KW
PS	PS	KW
PS	PS	KW
PS	PS	Synthetic.
DR	DR	OS
N-PSDB	N-PSDB	OS
AAH42878	AAH42878	Human immunodeficiency virus 1.
XX	XX	XX
PT	PT	PN WO2003011334-A1.
PT	PT	XX
PT	PT	PD 13-FEB-2003.
XX	XX	XX
PS	PS	PP 26-JUL-2002; 2002WO-EP008343.
XX	XX	PP 27-JUL-2001; 2001GB-00018367.
CC	CC	PR
CC	CC	XX
CC	CC	PA GLAXOSMITHKLINE BIOLOGICALS SA.
CC	CC	PA (GLAXO) GLAXO GROUP LTD.
CC	CC	PA
CC	CC	PI Ertel PF, Tite JP, Van Wely CA, Voss G;
CC	CC	XX
CC	CC	DR WPI; 2003-239474/23.
CC	CC	DR N-PSDB; ADE67692.
CC	CC	XX
CC	CC	PP Use of an HIV Tat, Nef or gp120 protein or polynucleotide in the manufacture of a prime-boost delivery for the prophylactic or therapeutic immunization of humans against HIV.
CC	CC	PT
CC	CC	PT manufacturing a vaccine for a prime-boost delivery for the prophylactic or therapeutic immunization of humans against HIV.
CC	CC	PR
CC	CC	PT
CC	CC	PT or therapeutic immunization of humans against HIV.
XX	XX	XX
PS	PS	PS Disclosure; Fig 1; 108pp; English.
XX	XX	XX
CC	CC	The present invention describes the use of an HIV Tat protein or polynucleotide, an HIV Nef protein or polynucleotide, or an HIV Tat protein or polynucleotide linked to an HIV Nef protein or polynucleotide or an HIV gp120 protein or polynucleotide in manufacturing a vaccine or therapeutic immunisation of humans against HIV. The protein or polynucleotide is delivered via a bombardment approach. Also described: (1) a recombinant DNA molecule comprising a Nef and/or Tat and/or gp120 gene in a vector which the gene of interest is inserted 3', to an enhanced HCMV IE1 promoter; (2) comprising a Nef and/or Tat and/or gp120 gene in a vector; and (3) a kit comprising at least two different vaccine compositions including: (a) a composition comprising particles coated with DNA encoding gp120 and nef and/or tat or nef/tat; and (b) a composition comprising gp120 and nef and/or tat or nef-tat DNA or proteins, where the DNA or proteins are coated onto the particles. The HIV proteins or polynucleotides are used to produce vaccines having anti-HIV activity. The recombinant DNA molecule can be used in manufacturing an HIV vaccine for the prophylactic or therapeutic immunisation of humans. The present sequence represents HIV-1 Lipop-nef-His fusion amino acid sequence, which is used in an example from the present invention.
XX	XX	XX
SQ	SQ	Sequence 324 AA;
Query	Query	Score 24.5%; Score 606; DB 4; Length 324;
Match	Match	Best Local Similarity 37.3%; Pred. No. 4.7e-44;
Matches	Matches	34; Mismatches 118; Indels 130; Gaps 15;
Qy	Qy	19 CSSHSMANTQMSDKIIIAHRGASGYLPEHTLESKALAFQAQDYLEQDLAMTKDGRLL 78
Db	Db	1 CSSHSMANTQMSDKIIIAHRGASGYLPEHTLESKALAFQAQDYLEQDLAMTKDGRLL 60
Dy	Dy	79 VVHDHFLDGLTDVAKFPHHRKGRRYVYDFTFLIKEIQSLEMTNFETKDGKQHQVYPN 138
Db	Db	61 VVHDHFLDGLTDVAKFPHHRKGRRYVYDFTFLIKEIQSLEMTNFETKDGK----- 113
Dy	Dy	139 RFPFLWKSHPRHTPDEIEFQGLEKSTGKGVGYYPEIKAPWFHHONGKDIAETLKVLR 198
Db	Db	114 -----W----- SKSFGVVG-WPTVRERMRRAFAADVGAGA-SRDLE 147
Dy	Dy	199 KGYDKTDKDMVLYLQTDFNELKRIKTELLQMGMDLKLVOLIAYTDWKEQDKPKGYWV 258
Db	Db	148 KHGAISSNTAATN----- AACAWLQEQQEBEVGFPPV 179
Dy	Dy	259 NYNTDWMFKEGAMAEVVKYADGVGHWYMLVNEEKSKDNIYVTPLVKELAQYNYVEVHPY 318
Db	Db	180 T----- PQQPLRPMYKAQVDLISHFL---KEKGGLSLLHSQR----- 214
Dy	Dy	319 TVRKDALPBFETTDYQMYDALLNKSATGKSMGKAGYVA 374
Db	Db	215 --RQDIDWYHTO----- GTFPDRNQYTPGPGRY----- PLTFGMCYQLV 255
Dy	Dy	375 LVRSDYKLYNKNSNSNSTLKNLGEHHRARAMDGGKAGVALVRSDFKFYEDANGTRDHKKG 434
Db	Db	256 PVEFD-KVEANKGENTSLLHPVSH---GMDDPEREVLWFRDSL-----AF 300
Dy	Dy	435 RHTARTSRSYKFDNKR1DSTSIGHHHHH 464
Qy	Qy	Sequence 324 AA;
Qy	Qy	Query Match 24.5%; Score 606; DB 7; Length 324;
Qy	Qy	Best Local Similarity 37.3%; Pred. No. 4.7e-44;
Qy	Qy	Mismatches 168; Conservative 34; Mismatches 118; Indels 130; Gaps 19 CSSHSMANTQMSDKIIIAHRGASGYLPEHTLESKALAFQAQDYLEQDLAMTKDGRLL 1 CSSHSMANTQMSDKIIIAHRGASGYLPEHTLESKALAFQAQDYLEQDLAMTKDGRLL

Qy	79	VVIIHDPLDGLTDVAKKPFHRHKDGRYVVIDPFTLKBISQSLLEMTEFNFTDKGKQAVVYN	138	Best Local Similarity	37.2%	Pred. No.	2.9e-43;
Db	61	VVIIHDPLDGLTDVAKKPFHRHKDGRYVVIDPFTLKBISQSLLEMTEFNFTDKGK-----	113	Matches	167;	Conservative	34;
Qy	139	RFPFLWKSFRHTFEDIEFQGLEKSTKKGVIGPEIRKAIFWPHHONGKIDIAETLKVLR	198	Qy	20	SSHSNNMANANTQMSDKIILAHRGSGYLPBHTLESKALAFAQDQDYLEQDLAMTKDGRLV	79
Db	114	-----W-----SKSSVVG-WPTVREMRRAEPAAUGVGAASRDL	147	Db	4	SSHSNNMANANTQMSDKIILAHRGSGYLPBHTLESKALAFAQDQDYLEQDLAMTKDGRLV	63
Qy	199	KYGDKTKTDMYQLQTFDNEFLKRIKTELLPQMGDLKLQJAYTDWKETQEDPKGNNY	258	Qy	80	VIHDIFLDGLTDVAKKPFHRHKDGRYVYDFTURKEIQSLEMTEFNFTDKGKQAVVYN	139
Db	148	RHGAIITSSNTAATN-----AACAWLEAQEEBEGVGPV	179	Db	64	VIHDIFLDGLTDVAKKPFHRHKDGRYVYDFTILKEIQSLEMTEFNFTDKGKQAVVYN	115
Qy	259	NNYDMMFKPGAMAEVVYKAQYGPWMLYNKEESKPDNIVYTPLVKELAQYNVEVHPY	318	Qy	140	FPLWKSFRHTFEDIEFQGLEKSTGKKGVIGPEIRKAIFWPHHONGKIDIAETLKVLR	199
Db	180	T-----POVPLRPMTYKAADVLSHL---KEKGLEGUHSOR-----	214	Db	116	-----SKSSVVG-WPTVREMRRAEPAAUDGVGAASRDL	150
Qy	319	TVRDALPEFFDVNQMYDALLNKSATGVFTDF---PDGVEFLKGTKSMDGGKAGVA	374	Qy	200	YGYDKKTDWVYLQPFDFNELKRIKTELLQMGMDLKLVOLIAVTDWKEYEQDPKGYVYN	259
Db	215	-RQDILDLWYHTQ-----GYFPDWQNYTPGPGRVY---PLTFGMCYKLV	255	Db	151	HGATISNTAATN-----AACAWLEAQEEEFGFPVY	182
Qy	375	LVRSDYKLYNKNSNSNSTKLNGEHRARMGGKAGVALVRSDFKYEFDANGTDRDKKG	434	Qy	260	YNYDMFMFKPGAMAEVVYKAQDGVGPWYMLVNKEESKPDNIVYTPLVKELAQYNVEVHPY	319
Db	256	PVBDP-KVEEANKGENTSLHPVSLH---GMDDPEREVLWRFDSRL-----AP	300	Db	183	-----POVPLRPMTYKAADVLSHL---KEKGLEGUHSOR-----	216
Qy	435	RHTARTSRSDYKFDYDKRIDSSTSGHHHH	464	Db	320	VRKDALPEFFDVNQMYDALLNKSATGVFTDF---PDGVEFLKGTKSMDGGKAGVA	375
Db	301	HHVARELHPEY-FKNN---CTSGHHHH	324	Db	217	-RQDILDLWYHTQ-----GYFPDWQNYTPGPGRVY---PLTFGMCYKLV	258
AC	AAY02354	standard; protein; 326 AA.		Qy	376	VRSDYKLYNKNSNSNSTKLNGEHRARMGGKAGVALVRSDFKYEFDANGTDRDKKG	435
XX	XX			Db	259	VEPD-KVEEANKGENTSLHPVSLH---GMDDPEREVLWRFDSRL-----APH	303
DT	09-JUL-1999	(first entry)		Qy	436	HTARTSRSDYKFDYDKRIDSSTSGHHHH	464
XX	XX			Db	304	HVARBLHPEY-FKNN---CTSGHHHH	326
RESULT				Search completed: June 22, 2005, 11:37:45			
ID.	AAY02354			Job time : 166 secs			
XX	XX						
DE	A representative LipoD-Nef fusion protein.						
XX	HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein; vaccine; HIV infection; protein D.						
OS	Synthetic.						
XX	Human immunodeficiency virus 1.						
PN	W09916884.A1.						
XX							
PD	08-APR-1999.						
XX							
PF	17-SEP-1998;	98WO-EP006040.					
XX							
PR	26-SEP-1997;	97GB-00020585.					
XX							
PA	(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.						
XX							
PI	Bruck C, Godart SAG, Marchand M;						
XX							
DR	WPI; 1999-302282/25.						
DR	N-PSDB; AAX35690.						
XX							
PT	HIV Tat or Nef protein linked to a fusion partner.						
PS	Disclosure: Fig 2: 66pp; English.						
XX	The present sequence represents a fusion protein comprising LipoD-HIV-1						
CC	Nef. The protein is exemplifies the fusion proteins of the invention. The						
CC	specification also describes fusion proteins comprising HIV-1 Tat						
CC	protein. The fusion protein can be used in a vaccine to prevent HIV						
CC	infection.						
XX	Sequence 326 AA;	24.1%;	Score 597;	DB 2;	Length 326;		
SQ	Query Match						

US-09-543-681A-6003 ; Sequence 6003, Application US/09543681A
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709-1.002-001
; CURRENT APPLICATION NUMBER: US/09/543, 681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO: 6003
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; US-09-543-681A-6003

Query Match 49:2%; Score 1219.5; DB 4; Length 364;
Best Local Similarity 64.7%; Pred. No. 4.1e-105; Indels 5; Gaps 3;
Matches 229; Conservative 50; MisMatches 70; DelIns 5; Gaps 3;

Qy 8 LSLLAAGVLAGCSSHSSMANMTQMSDKLIIAHRGASGVLPEHTLESKALAFQAQADYLE 67
Db 12 LKPLVAGVLLTLIS--LSTIA--QAAASDVKVIAHGRASGVLPEHTLPAKALAYAQAQADYLE 67

Qy 68 QDLAMTQDKGRLVVTDHFLGDLTDAKKFPHRHKDGRYVYDFIPLKEIQSLEMENFET 127
Db 68 QDLVMTKDDNQLVVLHDHYLDRVFTDAERYPVAIRKDGRYVAIDFTLDEKLKGKTEGFDI 127

Qy - 128 KDGKQAOQVYPNRPFLWKSHPRIHTPEDEIETPTGLEKSTGRKVGYPEKAPWFHONGK 187
Db 128 VDGKRVQSYPRFRPMGKSDFKIITQEEIPIQLINKSTQDQDIGIYPEKAPWFHONGK 187

Qy 188 DIAAETLKVLLKQYKGYDKKTKDMVYQLOTFDFNLKRKTELLPQMGMQDLKLVQOLIAYTDWKE 247
Db - 188 DITKQKLEVLRKQYGTQKSDDNVLQFQSDPNDLKRKTELMFQMGMQDLKLVQOLIAYTDWNE 247

Qy 248 TQEKDQKGYWNTYNNYDMMFKGAMAERVKYADGVGFBGWMVLYNKEESKPDNIVYPLVKE 307
Db 248 TYEKOPQGTWNTYNNYDMMFKGAMKEIATYADGIGDYNMLV-EESTDPKITLTGMAAD 306

Qy 308 LAQYNVHYPVTVRQALPEFTTDVQMDALLNGATGVYPTDPDTGVBFKL 361
Db 307 ARANKLTIHPTFTVRVDKLPKYAKDGDOLYDIYNNQAGAEVYFTDPDGVBFKL 360

RESULT 3 US-09-485-885-1 ; Sequence 1, Application US/09485885
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Cabezon, Silvia, Teresa
; APPLICANT: Delisse, Anne-Marie Eva Fernande
; APPLICANT: Gerard, Catherine Marie Ghislaine
; APPLICANT: Lombardo-Bencheikh, Angela
; FILE REFERENCE: B45107
; CURRENT APPLICATION NUMBER: US/09/485, 885
; CURRENT FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: PCT/EP98/05285
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: GB 9717953.5
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ For Windows Version 3.0
; SEQ ID NO: 8
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Homo sapien
; US-09-485-885-8

Query Match 22.8%; Score 565; DB 3; Length 220;
Best Local Similarity 68.0%; Pred. No. 1.7e-44; Indels 8; Gaps 2;
Matches 117; Conservative 3; MisMatches 3; DelIns 4; Gaps 2;

Qy 20 SSHSSMANMTQMSDKLIIAHRGASGVLPEHTLESKALAFQAQADYLE 79
Db 4 SSHSSMANMTQMSDKLIIAHRGASGVLPEHTLESKALAFQAQADYLE 63

Qy 80 VHDHFLGDTDAKKFPHRHKDGRYVYDFIPLKEIQSLEMENFET 128
Db 64 VHDHFLGDTDAKKFPHRHKDGRYVYDFIPLKEIQSLEMENFETMAMHGDPTLHE 123

Qy 129 -----DGKQAOQVYPNRPFLWKSHPFRTHTF 152
Db 124 YMLDLOPETTDLYGQQLNDSSEREDIDPGQAEPRD----AHYNIVTP 170

RESULT 4 US-09-485-885-8 ; Sequence 8, Application US/09485885
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Cabezon, Silvia, Teresa
; APPLICANT: Delisse, Anne-Marie Eva Fernande
; APPLICANT: Gerard, Catherine Marie Ghislaine
; APPLICANT: Lombardo-Bencheikh, Angela
; FILE REFERENCE: B45107
; CURRENT APPLICATION NUMBER: US/09/485, 885
; CURRENT FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: PCT/EP98/05285
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: GB 9717953.5
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ For Windows Version 3.0
; SEQ ID NO: 8
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Homo sapien
; US-09-485-885-8

Query Match 22.8%; Score 565; DB 3; Length 220;
Best Local Similarity 68.0%; Pred. No. 1.7e-44; Indels 8; Gaps 2;
Matches 117; Conservative 3; MisMatches 3; DelIns 4; Gaps 2;

Qy 20 SSHSSMANMTQMSDKLIIAHRGASGVLPEHTLESKALAFQAQADYLE 79
Db 4 SSHSSMANMTQMSDKLIIAHRGASGVLPEHTLESKALAFQAQADYLE 63

Qy 80 VHDHFLGDTDAKKFPHRHKDGRYVYDFIPLKEIQSLEMENFET 128
Db 64 VHDHFLGDTDAKKFPHRHKDGRYVYDFIPLKEIQSLEMENFETMAMHGDPTLHE 123

Qy 129 -----DGKQAOQVYPNRPFLWKSHPFRTHTF 152
Db 124 YMLDLOPETTDLYGQQLNDSSEREDIDPGQAEPRD----AHYNIVTP 170

RESULT 5 US-09-485-885-21 ; Sequence 21, Application US/09485885
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Cabezon, Silvia, Teresa
; APPLICANT: Delisse, Anne-Marie Eva Fernande
; APPLICANT: Gerard, Catherine Marie Ghislaine
; APPLICANT: Lombardo-Bencheikh, Angela
; FILE REFERENCE: B45107
; CURRENT APPLICATION NUMBER: US/09/485, 885
; CURRENT FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: PCT/EP98/05285
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: GB 9717953.5
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ For Windows Version 3.0
; SEQ ID NO: 1
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Homo sapien
; US-09-485-885-1

CURRENT APPLICATION NUMBER: US/09/485,885
 CURRENT FILING DATE: 2000-02-18
 PRIOR APPLICATION NUMBER: PCT/EP98/05285
 PRIOR FILING DATE: 1998-08-17
 PRIORITY NUMBER: GB 9717953.5
 NUMBER OF SEQ ID NOS: 23
 SOFTWARE: FastSEQ for Windows Version 3.0
 SEQ ID NO 21
 TYPE: PRT
 ORGANISM: Homo sapien

Query Match Score 22.7%; DB 3; Length 278;
 Best Local Similarity 34.5%; Pred. No. 3.8e-44;
 Matches 157; Conservative 29; Mismatches 79; Indels 190; Gaps 13;

Qy 20 SSSHSSNANTQMSKDKTIIIAHRCAGYLPEHTLESKALAFQQADYLEQDLMTKDGRLV 79
 Db 4 SSSHSSNANTQMSKDKTIIIAHRCAGYLPEHTLESKALAFQQADYLEQDLMTKDGRLV 63

Query Match Score 563; DB 3; Length 278;
 Best Local Similarity 34.5%; Pred. No. 3.8e-44;
 Matches 157; Conservative 29; Mismatches 79; Indels 190; Gaps 13;

Qy 80 VHDDHFDGLTDVAKKEPHRHKDGRYYVDFTLKEIOSLEMENPETKQDKQAAQVYPNR 139
 Db 64 VHDDHFDGLTDVAKKEPHRHKDGRYYVDFTLKEIOSLEMENPETKQDKQAAQVYPNR 111

Query Match Score 559.5; DB 3; Length 273;
 Best Local Similarity 91.1%; Pred. No. 7.8e-44;
 Matches 112; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

Qy 140 FPLWKSHFRHTPFDTEPIGGLEKSTGKQGYIPEKAPWPHQKQDIAETLKVLK 199
 Db 112 - - - - - MARFEDPTR - - - - - REPKLP - - - - - DLCTELNTSLQ- 137

Query Match Score 559.5; DB 3; Length 273;
 Best Local Similarity 91.1%; Pred. No. 7.8e-44;
 Matches 112; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

Qy 200 YGYDKKTDMDVYLTQDFNELLKRKTEELLPGCMGDLKLKVLIAYTDKETQKDPKGQYWN 259
 Db 138 - - - - - DIETCTVICKTV - - - - - LELTEV - - - - - 155

Query Match Score 559.5; DB 3; Length 273;
 Best Local Similarity 91.1%; Pred. No. 7.8e-44;
 Matches 112; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

Qy 260 YNYDMMFKPGAMAEEVVKYADGCGPMYMMVNKEESKPDNIVYTPVKELAQNVEVHPYT 319
 Db 156 - - - - - FEFAFKD - - - - - LFVYYRDSI - - - - - PHAACHKCIDFYSRIRELHYSDSVYGD 200

Query Match Score 559.5; DB 3; Length 273;
 Best Local Similarity 91.1%; Pred. No. 7.8e-44;
 Matches 112; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

Qy 320 VRKDALPFFTDVQMDYALLNKGATGVTDFPDTCVFFLKGIKSMDDGKAGVALVRSD 379
 Db 201 LEK - - - - - LNTNG- - - - - 208

Query Match Score 559.5; DB 3; Length 273;
 Best Local Similarity 91.1%; Pred. No. 7.8e-44;
 Matches 112; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

Qy 380 YKLYN - - - - - KNSSSNSTLKNLGEHEHARAMDGGKAGVALYRSDYKFYEDANGTRD 430
 Db 209 - - - - - LYNNLRLCRQKPLNPABKLRLHNEKRREFHLAG- - - - - 243

Query Match Score 559.5; DB 3; Length 273;
 Best Local Similarity 91.1%; Pred. No. 7.8e-44;
 Matches 112; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

Qy 431 HKKGR - HTARTSRSDYKFDYNKRDSTSGHHHHH 464
 Db 244 HYRQOCHSCCNRQBRLQRRETQVSGHHHHHH 278

RESULT 6
 CURRENT APPLICATION NUMBER: US/09/485,885
 Sequence 19, Application US/09485885
 Patent No. 6342224
 GENERAL INFORMATION:
 APPLICANT: Bruck, Claudine
 APPLICANT: Cabezon, Silvia, Teresa
 APPLICANT: Delisse, Anne-Marie Eva Fernande
 APPLICANT: Gerard, Catherine Marie Ghislaine
 APPLICANT: Lombardo-Bancheikh, Angelia
 TITLE OF INVENTION: Vaccine
 FILE REFERENCE: B45107

CURRENT APPLICATION NUMBER: US/09/485,885
 CURRENT FILING DATE: 2000-02-18
 PRIOR APPLICATION NUMBER: PCT/EP98/05285
 PRIOR FILING DATE: 1998-08-17
 PRIORITY NUMBER: GB 9717953.5
 NUMBER OF SEQ ID NOS: 23
 SOFTWARE: FastSEQ for Windows Version 3.0
 SEQ ID NO 19
 LENGTH: 227

RESULT 8
 US/09-485-885-6
 Sequence 6, Application US/09485885
 Patent No. 6342224
 GENERAL INFORMATION:
 APPLICANT: Bruck, Claudine
 APPLICANT: Cabezon, Silvia, Teresa
 APPLICANT: Delisse, Anne-Marie Eva Fernande
 APPLICANT: Gerard, Catherine Marie Ghislaine
 APPLICANT: Lombardo-Bancheikh, Angelia

RESULT 10
 APPLICANT: Lombardo-Bencheikh, Angela
 FILE REFERENCE: B45107
 CURRENT APPLICATION NUMBER: US/09/485,885
 CURRENT FILING DATE: 2000-02-18
 PRIOR APPLICATION NUMBER: PCT/EP98/05285
 PRIOR FILING DATE: 1998-08-17
 PRIOR APPLICATION NUMBER: GB 9717953.5
 PRIOR FILING DATE: 1997-08-22
 NUMBER OF SEQ ID NOS: 23
 SOFTWARE: FastSEQ for Windows Version 3.0
 SEQ ID NO: 6
 LENGTH: 371
 TYPE: PRT
 ORGANISM: Homo sapien
 US-09-485-885-6

Query Match 22.6%; Score 559.5; DB 3; Length 371;
 Best Local Similarity 91.1%; Pred. No. 1.2e-43;
 Matches 112; Conservative 1; Mismatches 49; Indels 1; Gaps 1;
 SEQ ID NO: 23

Qy 20 SSHSSNNANTQMSKDKIIIAHRGASGYLPETTLESKALAAQADYLQEDLAMTKDGRIV 79
 Db 4 SSHSSNNANTQMSKDKIIIAHRGASGYLPETTLESKALAAQADYLQEDLAMTKDGRIV 63

Qy 80 VHDDHFLDGLTDVAKKFPHRKDKGYYVIDFTLKEIQSLEMENFETKDGKQ-AQVYPN 138
 Db 64 VHDDHFLDGLTDVAKKFPHRKDKGYYVIDFTLKEIQSLEMENFETMFDQERR 123

Qy 139 RFP 141
 Db • 124 xLP 126

RESULT 9
 US-09-485-885-16
 Sequence 1.6, Application US/09485885
 SEQ ID NO: 6342224

GENERAL INFORMATION:
 APPLICANT: Bruck, Claudine
 APPLICANT: Cabezon, Silvia, Teresa
 APPLICANT: Delisse, Anne-Marie Eva Fernande
 APPLICANT: Gerard, Catherine Marie Ghislaine
 APPLICANT: Lombardo-Bencheikh, Angela
 TITLE OF INVENTION: Vaccine
 FILE REFERENCE: B45107
 CURRENT APPLICATION NUMBER: US/09/485,885
 CURRENT FILING DATE: 2000-02-18
 PRIOR APPLICATION NUMBER: PCT/EP98/05285
 PRIOR FILING DATE: 1998-08-17
 PRIOR APPLICATION NUMBER: GB 9717953.5
 PRIOR FILING DATE: 1997-08-22
 NUMBER OF SEQ ID NOS: 23
 SOFTWARE: FastSEQ for Windows Version 3.0
 SEQ ID NO: 16
 LENGTH: 227
 TYPE: PRT
 ORGANISM: Homo sapien
 US-09-485-885-16

Query Match 22.5%; Score 558; DB 3; Length 227;
 Best Local Similarity 100.0%; Pred. No. 8.1e-44;
 Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 SSHSSNNANTQMSKDKIIIAHRGASGYLPETTLESKALAAQADYLQEDLAMTKDGRIV 79
 Db 4 SSHSSNNANTQMSKDKIIIAHRGASGYLPETTLESKALAAQADYLQEDLAMTKDGRIV 63

Qy 80 VHDDHFLDGLTDVAKKFPHRKDKGYYVIDFTLKEIQSLEMENFETKDGKQ-AQVYPN 138
 Db 64 VHDDHFLDGLTDVAKKFPHRKDKGYYVIDFTLKEIQSLEMENFET 111

RESULT 11
 US-09-328-352-5514
 Sequence 5514, Application US/09328352
 SEQ ID NO: 6562958

GENERAL INFORMATION:
 APPLICANT: Gary L. Breton, et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 CURRENT APPLICATION NUMBER: US/09/3228,352
 CURRENT FILING DATE: 1999-06-04
 NUMBER OF SEQ ID NOS: 8252
 SEQ ID NO: 5514
 LENGTH: 386
 TYPE: PRT
 ORGANISM: Acinetobacter baumannii
 US-09-328-352-5514

Query Match 15.4%; Score 381.5; DB 4; Length 386;
 Best Local Similarity 31.6%; Pred. No. 5e-27;
 Matches 125; Conservative 51; Mismatches 145; Indels 75; Gaps 16;

Qy 6 LAISLLAAGVLAGCSSH--SSNNANTQMSKDKIIIAHRGASGYLPETTLESKALAAQ 61
 Db 13 LCLSLIS--LVGMDDDKETTPTPTEQLPKTLVVGHRGASAIRPENTLASTYQKAIID 69

Qy 62 QADYLEQDLAMTKDGRIVYTHDFDGLTDVAKKFPHRK---DG---RYVFDFT 112
 Db 70 GADFEPDLVSTGQYLVARHENEIGGTINVSTLSQADRKTKNIDGVDTGNTEDFT 129

Qy 113 LKEIQSLEMET--NFEETKDGKQAGYYPNRFPLWKSHPFTHTFENETIEFTQGLESTGK 169
 Db 130 LSELQQLKARERIPFRPANTAYNDLP-----VPTLQGILAEANTYKTKGK 178

Qy 170 VGIYPEIKAP-WFHONGKDIAB--TLKVLKKGYDRTDMVYLOTFDFNLKRIKTEL 226
 Db 179 IGYIETKHPYFKQN--LAMEDTLKTLAKYKTRDIAFVYLQSPEVNLKDLKREL 235
 Qy 227 LPQMGMDLKLKVOLIAYTDWKEQDKDPKGYWNNYNDWMFKEGAMAEVVKYADGVP-G 284
 Db 236 - -DLHKTLKHAQIQLQYDTSKTSRPADEVSQGKTVADLATAQGLKDVAKYANGVGPSTKG 293
 Qy 285 WYMLVNKEBESKDPDNIVYPTPLVKELAQYNNVEHVPYTRD----- 323
 Db 294 YLTENNDGSYKT---STFSDAHTAGLNSGATGVTDPPGVE 358
 Qy 324 -ALPEFTDVNOMYDALLINKSGATGVTDPPGVE 358
 Db 350 GALKEF---EAY---FKAGYDGVTFDPALGRE 376

RESULT 12
 Sequence 27882, Application US/09252991A.
 Patent No. 65511795
 GENERAL INFORMATION:
 APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 TITLE OF INVENTION: ABRUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 107196.136
 CURRENT APPLICATION NUMBER: US/09/252,911A
 CURRENT FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 27882
 LENGTH: 503
 TYPE: PRT
 ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-27882

Query Match 15.4%; Score 381; DB 4; Length 503;
 Best Local Similarity 30.5%; Pred. No. 8. 4e-7; Matches 119; Conservative 64; Mismatches 143; Indels 64; Gaps 13;

Qy 7 AUSLLAARGVLAGCSSHSSNM-----ANTOM-----KSDKIIIAHGRASGSGYLPEHTL 52
 Db 133 ACIMLAQGLASAHADITGNCVRAIDWAGQGSVRHPAGHRSPLVTAHGRASGVPBHTL 192
 Qy 53 ESKALAPAQQADYLEQDQDAMITDGRLVVTHDHLGDTDVAK--KPHRHK---DG--- 104
 Db 193 GAYALAVMGMADYEPDVIMTRDGKLVARHNBELGTVQKVVDGVEL 252
 Qy 105 -RYVYDITFLKEIQLSLEMENFTDKDQKAQVYPNRPFLWKSFRHTFEDIEFQGLE 163
 Db 253 TGFSEDETLAELKTRAIERIPT-----IRPGNARL-DGTFEPIPLQETIDLVKSLQ 304
 Qy 164 KSTGKKVKGYIYPETKAPNPHONGKDIAAETKLKVLYKGYDKTDMVYLOTFDFNLKRIK 223
 Db 305 ISQORTIGLYPIKHGTHFORGLAMERPLVTLHRYGLGRAPYFIQSFVNNUKELK 364
 Qy 224 TELLPGMDLKLKVQIAYTDWKEQDKDPKGYWNNYNDWMFKEGAMAEVVKYADGVP-G 283
 Db 365 RL-----TGIRLVLQ-----YGSQPYDQQAAGGSUTYAMATAKLRQYARYAYGVGP 413
 Qy 284 GYMLVNKEES---KEDUNIYPTPLVKEQYDQQAAGGSUTYAMATAKLRQYARYAYGVGP 327
 Db 414 DKSYVIPDANGNLGP-----TRFVDAHAAGLKVHPTFRAENSFLPAERSADGNPQ 468
 Qy 328 FFTDVNOMYDALLINKSGATGVTDPPGVE 357
 Db 469 SRGDLAGEIRAYLD-AGTDLGJLPSDQPHAV 497

US-08-956-171E-5231
 Sequence 5231, Application US/08956171E
 Patent No. 6593114
 GENERAL INFORMATION:
 APPLICANT: Charles Kunsch
 Gil H. Choi
 Patrick S. Dillon
 Craig A. Rosen
 Steven C. Barash
 Michael R. Farnon
 TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 NUMBER OF SEQUENCES: 5255
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/956,171E
 FILING DATE: 20-Oct-1997
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/009,861
 FILING DATE: January 5, 1996
 APPLICATION NUMBER: 08/781,986
 FILING DATE: January 3, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Mark J. Hyman
 REGISTRATION NUMBER: 46,789
 REFERENCE/DOCKET NUMBER: PB248P1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (240) 314-1224
 TELEFAX: (301) 309-8439
 INFORMATION FOR SEQ ID NO: 5231:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 316 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: Linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 5231:
 US-08-956-171E-5231

Query Match 10.9%; Score 269.5; DB 4; Length 316;
 Best Local Similarity 27.3%; Pred. No. 1e-16; Mismatches 116; Indels 107; Gaps 16;
 Matches 103; Conservative 48; Mismatches 116; Indels 107; Gaps 16;
 '8 ISLLAARGVLAGCSSHSSNM-----TOMKSDKIIIAHGRASGSGYLPEHTLSEKALA 58
 30 ISVPTAG-----AEQTNQIANKPQIAQWHTNLNTERPTTAHGRASGGAPEHTFQYDKS 84
 Qy '9 FAQ-QADYLEQDQDAMITDGRLVVTHDHLGDTDVAK--KPHRHK---DG--- 117
 Db 85 HNELKASYIETDQQTKDGLVAMDETVARTT-----NGHGVEDYTDDELK 132
 Qy 118 SLEMENPFTDKGKQAVQ-YPNRFLWKSFRHTFEDIEFQGLEKSTGKKVGYPEI 176
 Db 133 QLDAQSWENKCKYKARASYKNA-----KVPTLDELE-----RGPNANYIET 177
 Qy 177 KAPWFFHQNQKDIAAETKLKVYDQFDFBLKRKTFTELLPQMGW 232
 Db 178 KSPDVY----PGMEBQLLASLKKHLLNNNKKLKHQ----- 226
 Qy 233 DLKLVOLIAYTDWKEQDKDPKGYWNNYNDWMFKEGAMAEVVKYADGVP-GWMLVNKE 292
 Db 227 -NKYPLVKEVLDKGELQQ-----ENDQRLKEIRSTAIGLGD----- 262

RESULT 14
 US-08-781-986A-5231
 ; Sequence 5231, Application US/08781986A
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 5255
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEES: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 ; COMPUTER: HP Vectra 486/33
 ; SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; CURRENT APPLICATION NUMBER: US/08/781,986A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Benson, Bob
 ; REGISTRATION NUMBER: 30,446
 ; APPLICATION NUMBER: PR248PP
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-0504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 5231:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 316 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-781-986A-5231

Query 293 ESKPDNIVVTPVKE---LAQVNVEVHPYTVRKDALPEFFTVDQMYDAL-LNKGATG 347
 Db 263 -----YDLTEQNTTHLKD LGFIVHPY---WNEADMRLRINKYGVDG 302
 Qy 348 VFTDFDTGVEFLK 361
 Db 303 VFTNFADKYKEVTK 316

Query 348 VFTDFDTGVEFLK 361
 Db 303 VFTNFADKYKEVTK 316

RESULT 15
 US-09-134-001C-3115
 ; Sequence 3115, Application US/09134001C
 ; PATENT NO. 6380310
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC-007
 ; CURRENT APPLICATION NUMBER: US/09/134, 001C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064, 964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055, 779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO 1115
 ; LENGTH: 353
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus epidermidis
 ; US-09-134-001C-3115

Query Match 5 TLALSLAAGVLAAGCSHS 10.7%; Score 265; DB 3; Length 353;
 Best Local Similarity 27.0%; Pred. No. 3 1e-16;
 Matches 108; Conservative 52; Mismatches 114; Indels 126; Gaps 17;

Query 5 TLALSLAAGVLAAGCSHS 10.7%; Score 265; DB 3; Length 353;
 Db 20 TGLASLSSPSIAETSNHSPAPNOSSEKS 10.7%; Score 265; DB 3; Length 353;
 Best Local Similarity 27.0%; Pred. No. 3 1e-16;
 Matches 108; Conservative 52; Mismatches 114; Indels 126; Gaps 17;

Query 37 -----1IAHRGASGYLPEHTLESKALFAQ-QADYLFBDLAMTKDGR 78
 Db 80 NDTNHWTWNLTGERFTTAHRGASGYAPEHTPEAYDKSHNELGASYIETDLQRTDGHL 139
 Query 79 VVIIHDFLDGLTDVAKKEPHRHKDGRYVVIDFTKETOSLEMTEFNFTKDGKQAVQYPN 138
 Db 140 VAMHDEKVNRRT-----NIGHGRVDQLTKEKLQLDAGSWNRKHPPEAK--N 184
 Query 139 RFPWKSHPPIHTPDETEFIQGKSTKVGTYPEIAPWPHQNGDIAETFLVVK 198
 Db 185 KY---KNAKVPVLPB-----ILNRYGRNANTYIETKSPDVTFGMKOL---LDTLD 229

Query 199 KGY---DKKTDWVYLOPFDENELKRIKTELLPQMGDLKLVQLIAYTDWKETQEKDPK 254
 Db 230 KHDLLTQSLKRGHMYIQSFSGSSLEKVKH----HMNANIPLTRLMNKPEFLKATNQDLK 284
 Query 255 GYWVNTNYDWMFKPQAMAEVVKYADGVGPGWYMLVNKEESKPDNIVVTPVKEQYVNV 314
 Db 285 -----NIKSYAIGVGFEYTDL-----NIKNTNRLKNU--GFL 314
 Query 315 VHPYTVRKDALPEFFTVDQMYDALNKGATGTVFTDFPD 354
 Db 315 VHPYTV-----DDENOMRN--LNQIGVDGVTFNFAD 343

Search completed: June 22, 2005, 11:42:18
 Job time : 45 secs

Query 177 KAPWPHQNGDIAETFLVVKY---GYDKKTDWVYLOPFDENELKRIKTELLPQMGDM 232
 Db 178 KSPDVY---PMEQEQLASLKKHLLNNNKLKGHMIQSFSDESLKTHRQ-----226

Db 120 GFDIENCKKVQYTPGRPMGSDFRVTHFEEIEFQGLNHSRGNIGITYPEIKAPWFHH 179
 Qy 184 QNGKDAETLKVKKYQYDGTKTDMYLQTFDNEIKRITTELLPOMGDLKVLQIAYT 243
 Db 180 QEGKDAIAKTLLEVKKYQYTKDKVYQLCQTDADELKRNKEPLQMGMLNLYQIAYT 239
 Qy 244 DWKETOQDKPCKQYWNWYNNDMFKPQGAMAEMVYKVAADGVGPWYMLVNKEERSKPDNIVTP 303
 Db 240 DWNETQKQDPSWNTNNDWFKPQGAMKQVAEYADGIGPYHML-EBSQPNKILRG 298
 Qy 304 LVKELAQYNYEVHAPYTVRKDALPEFFTDVNQMDYALINKSGATGVFTDFDTGVFEL 360
 Db 299 MVQDAAQDNKLVVHPYTVRSKDLPEPTDVNQLYDALYNKACVNGLFTDFDKAVKFL 355

RESULT 2
 US-10-473-127-1113
 ; Sequence 1113, Application US/10473127
 ; GENERAL INFORMATION:
 ; APPLICANT: Zycos Inc.
 ; TITLE OF INVENTION: TRANSLATIONAL PROFILING
 ; FILE REFERENCE: 08191-026W01
 ; CURRENT APPLICATION NUMBER: US/10/473,127
 ; CURRENT FILING DATE: 2003-09-26
 ; PRIOR APPLICATION NUMBER: 60/279,495
 ; PRIOR FILING DATE: 2001-03-28
 ; PRIOR APPLICATION NUMBER: 60/292,544
 ; PRIOR FILING DATE: 2001-05-21
 ; PRIOR APPLICATION NUMBER: 60/310,801
 ; PRIOR FILING DATE: 2001-08-08
 ; PRIOR APPLICATION NUMBER: 60/326,370
 ; PRIOR FILING DATE: 2001-10-01
 ; PRIOR APPLICATION NUMBER: 60/336,780
 ; PRIOR FILING DATE: 2001-12-04
 ; PRIOR APPLICATION NUMBER: 60/358,985
 ; PRIOR FILING DATE: 2002-02-20
 ; NUMBER OF SEQ ID NOS: 2041
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 1113
 ; LENGTH: 451
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-473-127-1113

Query Match 26.7%; Score 66.5; DB 16; Length 451;
 Best Local Similarity 37.5%; Pred. No. 3..3e-45;
 Matches 194; Conservative 51; Mismatches 152; Indels 121; Gaps 18;

Db 1 1MDPKTLALSLIAAGVLAGCSSHSSNNANTOMKSDKLIIIAHRGASGYLPETTLESALAPA 60
 Db 1MDPKTLALSLIAAGVLAGCSSHSSNNANTOMKSDKLIIIAHRGASGYLPETTLESALAPA 60

Qy 1 1MDPKTLALSLIAAGVLAGCSSHSSNNANTOMKSDKLIIIAHRGASGYLPETTLESALAPA 60
 Db 1MDPKTLALSLIAAGVLAGCSSHSSNNANTOMKSDKLIIIAHRGASGYLPETTLESALAPA 60

Qy 61 QOADYLEQDLAMTKGRLVVIHDHFLGTDVAKFPHHRKQGRYYVDFTLKEIQSLE 120
 Db 61 QOADYLEQDLAMTKGRLVVIHDHFLGTDVAKFPHHRKQGRYYVDFTLKEIQSLE 120

Qy 121 MTENFETKDGKQAOQYVPNRFLWKSHR1HTPDETEF1QLELKVGKVKVGYYPEIKAPW 180
 Db 121 MTENFETMDLBORSQ----HCKPBE---GLE ARGEALGLV-GAQAPA 159

Qy 181 FHHONGKGDIAETLKVKKYQYDGTKTDMYLQTFDNEIKRITTELLPOMGMDL-KLVQL 239
 Db 160 TEEQEAASSSTLVEV----TLGEVPAAESPPDPPSPQGASSLPFT 201

Qy 240 IAYTDWKETQE---KDPKGYWNWYNNDMYMFPKG---AMAEVV----KY----AD 279
 Db 202 MNYPWPSQSYDSSNQEEBGPSTPQDLESEQAAALSRAKVLVHFLLLCYRAREPVTKA 261

Qy 280 GVGPP---GWMILVNKEEKSQDPNIVTPPLKELAQVNVEWAPYTVRKDALPEFFTDVNQMY 336
 Db 262 MLGSVVGWNQWFFPVFVFSKASSSLQLVFGLMEVDPIGHLYI-----FATCLGLSY 313

RESULT 3
 US-10-687-060-7
 ; Sequence 7, Application US/10687060
 ; Publication No. US20050033022A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bruck, Claudine
 ; APPLICANT: Godart, Stephane Andre Georges
 ; APPLICANT: Marc-Han, Martine
 ; TITLE OF INVENTION: Fusion Proteins Comprising HIV-1 Tat
 ; FILE REFERENCE: B45110
 ; CURRENT APPLICATION NUMBER: US/10/687,060
 ; CURRENT FILING DATE: 2003-10-16
 ; PRIOR APPLICATION NUMBER: US/09/509,239
 ; PRIOR FILING DATE: 2000-03-23
 ; PRIOR APPLICATION NUMBER: PCT/EP98/06040
 ; PRIOR FILING DATE: 1998-09-17
 ; PRIOR APPLICATION NUMBER: GB 9720585.0
 ; PRIOR FILING DATE: 1997-09-26
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO: 7
 ; LENGTH: 144
 ; TYPE: PRT
 ; ORGANISM: Pichia pastoris
 US-10-687-060-7

Query Match 26.3%; Score 651; DB 17; Length 144;
 Best Local Similarity 100.0%; Pred. No. 4.6e-45;
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 1MDPKTLALSLIAAGVLAGCSSHSSNNANTOMKSDKLIIIAHRGASGYLPETTLESALAPA 60
 Db 1MDPKTLALSLIAAGVLAGCSSHSSNNANTOMKSDKLIIIAHRGASGYLPETTLESALAPA 60

Qy 61 QOADYLEQDLAMTKGRLVVIHDHFLGTDVAKFPHHRKQGRYYVDFTLKEIQSLE 120
 Db 61 QOADYLEQDLAMTKGRLVVIHDHFLGTDVAKFPHHRKQGRYYVDFTLKEIQSLE 120

Qy 121 MTENFET 127
 Db 121 MTENFET 127

RESULT 4
 US-10-203-013-15
 ; Sequence 15, Application US/10203013
 ; Publication No. US20030158134A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SmithKline Beecham Biologicals S.A.
 ; TITLE OF INVENTION: No. US 10/203,013
 ; FILE REFERENCE: B45209
 ; CURRENT APPLICATION NUMBER: US/10/203,013
 ; CURRENT FILING DATE: 2002-07-31
 ; PRIOR APPLICATION NUMBER: GB 00022004
 ; PRIOR FILING DATE: 2000-01-31
 ; PRIOR APPLICATION NUMBER: GB 0009336.9
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: GB 0013806.5

/ PRIOR FILING DATE: 2000-06-06

/ PRIOR APPLICATION NUMBER: PCT/EP00/05998

/ PRIOR FILING DATE: 2000-06-28

/ NUMBER OF SEQ ID NOS: 31

/ SOFTWARE: FastSEQ for Windows Version 3.0

/ SEQ ID NO: 15

/ LENGTH: 324

/ TYPE: PRT

/ ORGANISM: Homo sapiens

/ US-10-203-03-15

US-10-687-060-15

Query Match

24 5% ; Score 606 ; DB 17 ; Length 324 ;

Best Local Similarity 37.3% ; Pred. No. 6.9e-41 ;

Matches 168 ; Conservative 34 ; Mismatches 118 ; Indels 130 ; Gaps 15 ;

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

/

Qy 79 VVIIDHFLGLTDVAKKPFPHRKDRYVYIDFTLKEIQSLEMTENPETDKQAOQVYN 138
 Db 61 VVIIDHFLGLTDVAKKPFPHRKDRYVYIDFTLKEIQSLEMTENPETGGK----- 113

Qy 139 RFLPKSHFRHTFEDIEFTQGLEKSTKXVGIYPEIKAQWPHHQNGKDIAAETLKVLK 198
 Db 114 -----W-----SKSSVVG-WPTVREMRRAEPAAQVGASAERDLK 147

Qy 199 KGYDKKTMVYLQTFDNEELKRKITELLPQMGMLKLKVQIAYTDKETQEPDKGYW 258
 Db 148 KHGATTSNTAATN-----AACAWLQAQEEEVGFPV 179

Qy 259 NYNDWNNPKGAMAEVKAYDGVGPWYMLVNKEESKPDNIVTPLYKELAQYINVHPY 318
 Db 180 T-----PQVPLRMTYKAVDLSHF-----KEKGLELHSOR----- 214

Qy 319 TVRKDALPEFTDVNQMYDALLNKSGATGVFTDF-----PDTGVEFLKGTKSMDGKGAGVY 374
 Db 215 -----RQDLDLWLYHTQ-----GVPDWQNYTPGPGVRY-----PLTFGWCYKLV 255

Qy 375 LYRSDYKLYNKNSSNSSLKLUGEHARRAAMDGGKAGVALVRSDFKYEADNGTRDHKG 434
 Db 256 PVEPD-KVEEANKGENTSLHPVSLH-----GMDPEREVLBWRFDRL-----AP 300

Qy 435 RHTARTSRSDYKFDYDNKRIOTSQGHHHH 464
 Db 301 HHVARELHPEY-FKN-----CTSGHHHHH 324

RESULT 7
 US-10-203-013-19
 ; Sequence 19, Application US/10203013
 ; Publication No. US2003015814A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SmithKline Beecham Biologicals S.A.
 ; TITLE OF INVENTION: No. US20030158134A1el Use
 ; FILE REFERENCE: B45209
 ; CURRENT APPLICATION NUMBER: US/10/203,013
 ; CURRENT FILING DATE: 2002-07-31
 ; PRIOR APPLICATION NUMBER: GB 0002200-4
 ; PRIOR FILING DATE: 2000-01-31
 ; PRIOR APPLICATION NUMBER: GB 00093336-9
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: GB 0013806-5
 ; PRIOR FILING DATE: 2000-06-06
 ; PRIOR APPLICATION NUMBER: PCT/EP00/05998
 ; PRIOR FILING DATE: 2000-06-28
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO: 19
 ; LENGTH: 326
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-203-013-19

Query Match 24.1%; Score 597; DB 14; Length 326;
 Best Local Similarity 37.2%; Pred. No. 3.8e-40; Indels 130; Gaps 15;
 Matches 167; Conservative 34; Mismatches 118; Del 130; Gap 15;

Qy 20 SSHSSNMANTONMSKDKIIIAHRGASGYLPEHTLESKALAAQOQDYLEQDLAMTKDGLV 79
 Db 4 SSHSSNMANTONMSKDKIIIAHRGASGYLPEHTLESKALAAQOQDYLEQDLAMTKDGLV 63

Qy 80 VTHDHFGLDGLTDVAKKPFPHRKDRYVYIDFTLKEIQSLEMTENPETDKQAOQVYN 139
 Db 64 VTHDHFGLDGLTDVAKKPFPHRKDRYVYIDFTLKEIQSLEMTENPETDKQAOQVYN 115

Qy 140 FPLWKSHPRIHTFEDIEFTQGLEKSTGKKGIVYIPEIKAQWPHHQNGKDIAAETLKVLK 199
 Db 116 -----W-----SKSSVVG-WPTVREMRRAEPAAQVGVAASRDLEK 150

Qy 200 YGDICKTDMYLOPQDFNEELKRKITELLPQMGMLKLVQIAYTDKETQEPDKGYW 259
 Db 151 HGATTSNTAATN-----AACAWLQAQEEEVGFPV 182

Qy 260 YNDWNNPKGAMAEVKAYDGVGPWYMLVNKEESKPDNIVTPLYKELAQYINVHPY 319
 Db 183 -----PQVPLRMTYKAVDLSHF-----KEKGLELHSOR----- 216

Qy 320 VRKDALPEFTDVNQMYDALLNKSGATGVFTDF-----PDTGVEFLKGTKSMDGKGAGVY 375
 Db 217 -----W-----SKSSVVG-WPTVREMRRAEPAAQVGVAASRDLEK 150
 Qy 200 YGDICKTDMVYLQTFDNEELKRKITELLPQMGMLKLVQIAYTDKETQEPDKGYW 259
 Db 116 -----W-----SKSSVVG-WPTVREMRRAEPAAQVGVAASRDLEK 150

Qy 376 VRSVDYKLYNKNSSSNSTLKNLGEHHRARAMDGGKAGYALVRSVDYKFYEDANGTRDHKKGR 435
 Db 259 VEPD-KVEANKGENTSLHPVSLH---GMDDPEREVLWRFDSRL-----AFH 303

Qy 436 HTPARTSRSDYKFYDNKRILDSTSCHHHHH 464
 Db 304 HWARELHPEY---FKN---CTSGHHHHHH 326

RESULT 9
 US-10-485-048-19
 ; Sequence 19, Application US/10485048
 ; Publication No. US2005005865781
 ; GENERAL INFORMATION
 ; APPLICANT: ERTL, Peter Franz
 ; TITE: John Phillip
 ; APPLICANT: VAN WELY, Catherine Anne
 ; APPLICANT: VOSS, Gerald
 ; TITLE OF INVENTION: Vaccine Comprising GP120 and NEF and/or
 ; TITLE OF INVENTION: TAT for the Immunisation Against HIV
 ; FILE REFERENCE: B45279
 ; CURRENT APPLICATION NUMBER: US/10/485,048
 ; CURRENT FILING DATE: 2004-01-26
 ; PRIOR APPLICATION NUMBER: PCT/EP02/08343
 ; PRIOR FILING DATE: 2002-07-26
 ; PRIOR APPLICATION NUMBER: GB 0118367.2
 ; PRIOR FILING DATE: 2001-07-27
 ; NUMBER OF SEQ ID NOS: 98
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 19
 ; LENGTH: 326
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-485-048-19

Query Match 24.1%; Score 597; DB 17; Length 326;
 Best Local Similarity 37.2%; Pred. No. 3. 8e-10;
 Matches 167; Conservative 34; Mismatches 118; Indels 130; Gaps 15;

Qy 20 SSSHSNNANTQMSKDKIIIAHRSAGSGLPEHTLESKALAAQADLEQDLMTKDGRV 79
 Db 4 SSSHSNNANTQMSKDKIIIAHRSAGSGLPEHTLESKALAAQADLEQDLMTKDGRV 63

Qy 80 VTHDHFELDGLTDVAKKEPKPHRKDRGYYVIFTLKRISLMTENFETDKQKAQVYPNR 139
 Db 64 VTHDHFELDGLTDVAKKEPKPHRKDRGYYVIFTLKRISLMTENFETDKGK-----115

Qy 140 FPLWKSHPRHTFDETEFIQLEKSTGKCVGYPEKAPMFKHONGKIDAAETLKVKK 199
 Db 116 ---W-----SKSSVVG-WPTVTRMRRAEPADGVGAARDLKK 150

Qy 200 YGYDKKTDMVYLQTFDFNELKRKITEBLQKETQKDPKGWYN 259
 Db 151 HGAITSNSNATN-----AACAWLEAQEEEVGFPTT 182

Qy 260 YNYDMFKPGAMAEVVKYADCYGPWYMLVNKEESKPDNIVYTPLYKELQYNNVEVHPY 319
 Db 183 -----PQVPLRPMTYKAVDLSHFL---KEKGSLGLJHSQR-----216

Qy 320 VRKDALPEFFDVMQNYDALLNKSATGVFTDF---PDTGVEFLKGKSMGDKKAGVAL 375
 Db 217 -RQDILWLYHTQ-----GYFPDWQNTGPQFVRY---PLTFGWCYKLV 258

Qy 376 VRSVDYKLYNKNSSSNSTLKNLGEHHRARAMDGGKAGYALVRSVDYKFYEDANGTRDHKKGR 435
 Db 259 VEPD-KVEANKGENTSLHPVSLH---GMDDPEREVLWRFDSRL-----AFH 303

Qy 436 HTPARTSRSDYKFYDNKRILDSTSCHHHHH 464
 Db 304 HWARELHPEY---FKN---CTSGHHHHHH 326

RESULT 10
 US-10-03-013-17
 ; Sequence 17, Application US/10203013
 ; Publication No. US20030158134A1
 ; GENERAL INFORMATION
 ; APPLICANT: SmithKline Beecham Biologicals S.A.
 ; TITLE OF INVENTION: No. US20030158134A1el Use
 ; FILE REFERENCE: B45209
 ; CURRENT APPLICATION NUMBER: US/10/203,013
 ; CURRENT FILING DATE: 2002-07-31
 ; PRIOR APPLICATION NUMBER: GB 0002200-4
 ; PRIOR FILING DATE: 2000-01-31
 ; PRIOR APPLICATION NUMBER: GB 00093336-9
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: GB 0013805-5
 ; PRIOR FILING DATE: 2000-05-06
 ; PRIOR APPLICATION NUMBER: PCT/EP00/05998
 ; PRIOR FILING DATE: 2000-05-28
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO: 17
 ; LENGTH: 411
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-203-013-17

Query Match 24.0%; Score 594.5; DB 14; Length 411;
 Best Local Similarity 35.6%; Pred. No. 8.4e-40;
 Matches 178; Conservative 45; Mismatches 134; Indels 143; Gaps 19;

Qy 19 CSSHSNNANTQMSKDKIIIAHRSAGSGLPEHTLESKALAAQADLEQDLMTKDGRV 78
 Db 1 CSSHSNNANTQMSKDKIIIAHRSAGSGLPEHTLESKALAAQADLEQDLMTKDGRV 60

Qy 79 VVTHDHFELDGLTDVAKKEPKPHRKDRGYYVIFTLKRISLMTENFETDKQKAQVYPN 138
 Db 61 VVTHDHFELDGLTDVAKKEPKPHRKDRGYYVIFTLKRISLMTENFETDKGK-----113

Qy 139 RFLPKNSHFR1LTFDDEBIEFQLEKSTGKCVGYPEKAPMFKHONGKIDAAETLKVKK 198
 Db 114 ---W-----SKSSVVG-WPTVTRMRRAEPADGVGAARDL 147

Qy 199 KYGYDKKTD-----MVYLQTFDFNELKRKITEBLQKEMGMDLK-LVQLJAYTDWK-----246
 Db 148 KHGATTSNTAAATNACAWLIAQBEBVGPVPTPVPLRPMTYKAVDLSHFLKGGLB 207

Qy 247 ---ETQEKOPKGTVW---NNYDWMMPKGAMAEVVKYADGVGP-----GW-YMLVN- 290
 Db 208 GLIHQRQRDILWLYHTQGYFPWQ-----NVTPGPGTVRYPLTFGWCYKLV 257

Qy 291 -----KEEKSPDNTIVYTPLYKELA-----QVNVEVPTVYTKDABBFPTD 331

Db 258 EPDKYEEANKEN---TSLLHPVSLHGMDDPEREVLEWRFDSRLPMPHYARELHBYPKN 314

Qy 332 VNQMDLNLNSKGATGVFTDFP-DTGVBFPLKGKSMGDKKAGVALVRSVDYKLYNNNSSN 390
 Db 315 C-----TSBVPDRLPMPKHPGSQP-----KTACTN 340

Qy 391 STLNKLGHEH----HRADAMDGSKAGVALVRSVDYKFYEDANGTRDHKKGRHTARTSRSD 444
 Db 341 CYCKCCPFHQVCPFTKALGSIYGR-----KKRQRQRREPPQGSATHOVSLSKQTSQSR 394

Qy 445 YKFYDNKR1DSTSGBHHHHH 464
 Db 395 ---GPPTGPKETSGHHHHHH 411

RESULT 11
 US-10-68-060-17
 ; Sequence 17, Application US/10203013
 ; Publication No. US20050031022A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bruck, Claudine

PRIOR FILING DATE: 2000-06-28
 NUMBER OF SEQ ID NOS: 31
 SOFTWARE: FastSEQ for Windows Version 3.0
 SEQ ID NO 21
 LENGTH: 413
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-203-013-21

Query Match 23.6%; Score 585.5; DB 17; Length 413;
 Best Local Similarity 35.5%; Pred. No. 4.6e-39;
 Matches 177; Conservative 45; Mismatches 134; Indels 143; Gaps 19;

Query 20 SSSHSSNANTOMKSDDKIIIAHRGASGYLPENTLESKALAFAAQQADYLEQDLMATKDGRLV 79
 SEQ ID NO 21
 LENGTH: 413
 TYPE: PRT
 ORGANISM: Pichia pastoris
 US-10-687-060-21

Query Match 23.6%; Score 585.5; DB 17; Length 413;
 Best Local Similarity 35.5%; Pred. No. 4.6e-39;
 Matches 177; Conservative 45; Mismatches 134; Indels 143; Gaps 19;

Query 20 SSSHSSNANTOMKSDDKIIIAHRGASGYLPENTLESKALAFAAQQADYLEQDLMATKDGRLV 79
 Db 4 SSSHSSNANTOMKSDDKIIIAHRGASGYLPENTLESKALAFAAQQADYLEQDLMATKDGRLV 63

Query 20 SSSHSSNANTOMKSDDKIIIAHRGASGYLPENTLESKALAFAAQQADYLEQDLMATKDGRLV 79
 Db 4 SSSHSSNANTOMKSDDKIIIAHRGASGYLPENTLESKALAFAAQQADYLEQDLMATKDGRLV 63

Db 4 VHDHFELDGLTDVAKCPHHRKDGRRYVYIDFTLKEIQTOSLENTENPETDGKQAOQVYPNR 139
 Db 4 VHDHFELDGLTDVAKCPHHRKDGRRYVYIDFTLKEIQTOSLENTENPETDGKQAOQVYPNR 139

Db 116 ---W-----SKSSVVG-WPTVRERMRRAEPADGGAASRDLEK 150

Db 116 ---W-----SKSSVVG-WPTVRERMRRAEPADGGAASRDLEK 150

Db 200 YGYDKKTD-----MVYLQTPDFNEBLKRKIKTELLPQMGMDLK-LVOLIAYTDWK---- 246
 Db 151 HGAITSSNTAAACAWLAQEEBEVGFPTPQPLRPMTYKAAVDLSHFLKEKGLEG 210

Db 116 ---W-----SKSSVVG-WPTVRERMRRAEPADGGAASRDLEK 150

Db 247 ---ETOBKDPKGYWW----NNYNDMFKPGAMAETVYKAADCVGVP----GW-YMLVN-- 290
 Db 211 LIHSQRQDIDLWVHTHQGYPPDQ----NYTPGPVTPYPLTFCWYKLVPVE 260

Db 151 HGAITSSNTAAACAWLAQEEBEVGFPTPQPLRPMTYKAADLSPHFLKEKGLEG 210

Db 247 ---ETOBKDPKGYWW----NNYNDMFKPGAMAETVYKAADCVGVP----GW-YMLVN-- 290

Db 211 LIHSQRQDIDLWVHTHQGYPPDQ----NYTPGPVTPYPLTFCWYKLVPVE 260

Db 318 -----TSEVPDPRLPEWKHPGSQL-----QYNYEVHPYTPVLRKELA----- 332
 Db 261 PDKVEEANGKEN---TSLLHPVSLHGMDDPEREVLWFRFDSRLAFHVARELHPETFKNC 317

Db 333 NQMYDALLNKGATGVTDFP-DTGYBFPLKGTKSMMDGGKAGVALYRSDDLYKLNKSSNS 391
 Db 318 -----TSEPYDPRLPEWKHPGSQL-----HRAMDGKGAGVALYRSDDLYKLNKSSNS 391

Db 392 TLKLNGBH-----HRAMDGKGAGVALYRSDDLYKLNKSSNS 391
 Db 344 YCKKCCPFHCQVCPITAKLGSYGR-----KKRQRERRPPQGSQTHQVSLSKQPTQSQR- 396

Db 446 KFYDNKRIDSTSGHHHHHHH 464
 Db 397 -GDPtgPKETSGHHHHHH 413

RESULT 15
 US-10-485-048-21
 Sequence 21, Application US/10485048
 Publication No. US2005005857A1
 GENERAL INFORMATION:
 APPLICANT: ERTL, Peter Franz
 APPLICANT: VAN WELY, Catherine Anne
 APPLICANT: VOSS, Gerald
 TITLE OF INVENTION: Vaccine Comprising GP120 and NEP and/or FILE REFERENCE: B45279
 FILE REFERENCE: B45279
 CURRENT APPLICATION NUMBER: US/10/485,048
 CURRENT FILING DATE: 2004-01-26
 PRIOR APPLICATION NUMBER: PCT/EP02/08343
 PRIOR FILING DATE: 2002-07-26
 PRIOR APPLICATION NUMBER: GB 0118337.2
 PRIOR FILING DATE: 2001-07-27
 NUMBER OF SEQ ID NOS: 98
 SEQ ID NO 21
 LENGTH: 413
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-485-048-21

Query Match 23.6%; Score 585.5; DB 17; Length 413;
 Best Local Similarity 35.5%; Pred. No. 4.6e-39;

RESULT 14
 US-10-687-060-21
 Sequence 21, Application US/10687060
 Publication No. US20050033022A1
 GENERAL INFORMATION:
 APPLICANT: Bruck, Claudine
 APPLICANT: Godard, Stephane Andre Georges
 APPLICANT: Marc-Han, Martine
 TITLE OF INVENTION: Fusion Proteins Comprising HIV-1 TAT
 TITLE OF INVENTION: and/or Nef Proteins
 FILE REFERENCE: B45110
 CURRENT APPLICATION NUMBER: US/10/687,060
 CURRENT FILING DATE: 2003-10-16
 PRIOR APPLICATION NUMBER: US/09/509,239
 PRIOR FILING DATE: 2000-03-23
 PRIOR APPLICATION NUMBER: PCT/EP08/06040
 PRIOR FILING DATE: 1998-09-17
 PRIOR APPLICATION NUMBER: GB 9720585.0
 PRIOR FILING DATE: 1997-09-26
 NUMBER OF SEQ ID NOS: 27
 SOFTWARE: FastSEQ for Windows Version 3.0
 SEQ ID NO 21
 LENGTH: 413

	Matches	177;	Conservative	45;	Mismatches	134;	Indels	143;	Gaps	19;
Qy	20	SSHSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAAQADYLEODLAMTKDGRIV	79							
Db	4	SSHSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAAQADYLEODLAMTKDGRIV	63							
Qy	80	VIHDHFLGLIDVAKKEPHHRKGRGYVDFTLKEISLEMTEFETDQKAOVYPNR	139							
Db	64	VIHDHFLGLIDVAKKEPHHRKGRGYVDFTLKEISLEMTEFETDQKAOVYPNR	115							
Qy	140	FPLWKSHFRHTFDEIEFIGLEKSTGKVGTYPEIKAPMFHHONGKODIAETTLVLUCK	199							
Db	116	W-----SKSUVG-WPTVTRMRRAFDGVAASRDLEK	150							
Qy	200	YGYDKKTD-----MVLQLQTFDENFLKRKTELLPQMGNDLK-MVQLTAYTDWK-----	246							
Db	151	HGAITSNTAATNAACAWLAEQEEBVGFPPTPQVBLRPTYKAADVLSHFLKEGGGLBG	210							
Qy	247	--ETQEKDPKGYW--NNTYNDMFKPGAMAEVVKYANGVGP-----GW-YMLVN--	290							
Db	211	LHHSQRQDLDLWYHTQGYFPDNO-----NYPTRPGVRYPLTGFWCYKLVPVE	260							
Qy	291	---KBEISKPDNIVTPVKELA-----QYNEVEHYYTVRKDALPFFTDV	332							
Db	261	PDKVEEANKGEN--TSLLHPVSLHGMDPDERELEWRFDSSLRAPHVARELHPFYKNC	317							
Qy	333	NOMYDALLNKGATGVTDFP-DTGVBELKQIKSMDDGKAGVALYRSDYKLYNKNSNS	391							
Db	318	-----TSEPVDPDRLEPWRKPGSQP-----KFACTNC	343							
Qy	392	TIKNLGEH-----HRARAMDGKAGVALYRSDYKFEYEDANGTRDHKKGRHTARTSRSDY	445							
Db	344	YCKKCCFHCOVCFITKALGSIYGR-----KKRQRPRPQGSOTHOVSLSKOPPSQSR-	396							
Qy	446	KFYDNKRIDSSTSGHHHHHH	464							
Db	397	-GDPTRGPKETSGHHHHHH	413							

Search completed: June 22, 2005, 11:54:14
 Job time : 161 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using bw model

Run on: June 22, 2005, 11:31:10 ; Search time 41 Seconds

(without alignments)
 1088.892 Million cell updates/sec

Title: US-09-719-379A-81

Perfect score: 2477

Sequence: 1 MDPTKLALSLLAAGVLAGGS.....YKFYDNKRIDSTSGHHHHHH 464

Scoring table: BL05162

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_79:*

1: Pir1:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

RESULT 1
 A43576
 Glycerophosphodiester phosphodiesterase (EC 3.1.4.46) precursor - Haemophilus influenzae
 N;Alternative names: immunoglobulin D-binding protein; protein D
 C;Species: Haemophilus influenzae
 A;Variety: isolate 772
 C;Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 08-Oct-1999
 C;Accession: A43576
 R;Janson, H.; Heden, L.O.; Grubb, A.; Ruan, M.; Forsgren, A.
 A;Title: Protein D, an immunoglobulin D-binding protein of Haemophilus influenzae: clonal
 A;Reference number: A43576; PMID:9109998; PMID:1387023
 A;Molecule type: DNA
 A;Residues: 1-364 <JAN>
 A;Cross-references: GB: M37487; NID: 9148970; PIDN: AAA24998.1; PID: 9148971
 A;Experimental source: isolate 772
 C;Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen
 F;1-18/Domain: signal sequence #status predicted <SIG>
 F;19-364/Product: glycerophosphodiester phosphodiesterase #status predicted <MAT>

Result No.	Score	Query	Match	Length	DB ID	Description
1	1920	glycerophosphodiester	1916	77.5	364 2 A43576	glycerophosphodiester
2	1916	glycerophosphodiester	1914	77.4	364 2 S59332	glycerophosphodiester
3	1914	glycerophosphodiester	1910	77.3	364 2 S59331	glycerophosphodiester
4	1910	glycerophosphodiester	1902	77.1	364 2 S59334	glycerophosphodiester
5	1902	glycerophosphodiester	1885	76.8	364 2 G64086	glycerophosphodiester
6	1885	glycerophosphodiester	1261	76.1	364 2 S59333	glycerophosphodiester
7	1261.5	glycerophosphodiester	1261.5	50.9	358 2 F85863	glycerophosphodiester
8	1261.5	glycerophosphodiester	1255	50.5	358 2 S15245	glycerophosphodiester
9	1255	glycerophosphodiester	1233	49.8	356 2 AH0791	glycerophosphodiester
10	1255	glycerophosphodiester	1210	48.8	371 2 AC0466	glycerophosphodiester
11	1210	glycerophosphodiester	1074	43.4	356 2 F71346	probable Glyceroph
12	1074	glycerophosphodiester	957	38.6	371 2 H82497	glycerophosphodiester
13	957	glycerophosphodiester	425	17.2	1027 2 AC1841	glycerophosphodiester
14	425	glycerophosphodiester	415.5	16.8	786 2 D75630	glycerophosphodiester
15	415.5	glycerophosphodiester	383	15.5	383 2 S83603	glycerophosphodiester
16	383	glycerophosphodiester	368	14.9	356 2 P87554	glycerophosphodiester
17	368	glycerophosphodiester	354	14.3	389 2 AH1937	glycerophosphodiester
18	354	glycerophosphodiester	353	14.3	372 2 T45228	glycerophosphodiester
19	353	glycerophosphodiester	347.5	14.0	392 2 D95770	hypothetical prote
20	347.5	glycerophosphodiester	321.5	13.0	293 2 I40418	glycerophosphodiester
21	321.5	glycerophosphodiester	269.5	10.9	309 2 H89662	glycerophosphodiester
22	269.5	glycerophosphodiester	262.5	10.6	284 2 DB4112	probable Glyceroph
23	262.5	glycerophosphodiester	246.5	10.0	375 2 A83352	glycerophosphodiester
24	246.5	glycerophosphodiester	222	9.0	243 2 E89927	glycerophosphodiester
25	222	glycerophosphodiester	191	7.3	247 2 E89956	hypothetical prote
26	191	glycerophosphodiester	179.5	7.2	239 2 F69361	glycerophosphodiester
27	179.5	glycerophosphodiester	172.5	7.0	583 2 AH1151	glycerophosphodiester

ALIGNMENTS

30	171	6.9	249 2 G96952	glycerophosphoryl
31	170.5	6.9	583 2 A11510	C-terminal domain
32	168	6.8	232 2 AD1236	glycerophosphodiest
33	167	6.7	235 2 A11598	glycerophosphoryl
34	165.5	6.7	285 2 E75317	hypothetical prote
35	156.5	6.3	240 2 G90544	outer membrane pro
36	154	6.2	353 2 C64187	agrocinopine phosph
37	152.5	6.2	268 2 AP3245	hypothetical prote
38	151.5	6.1	763 2 F96693	hypothetical prote
39	148.5	6.0	271 2 B86918	hypothetical prote
40	147.5	6.0	241 2 S73747	glycerophosphoryl
41	146	5.9	770 2 T04792	hypothetical prote
42	145	5.9	587 2 G89785	hypothetical prote
43	144.5	5.8	244 2 D64232	glycerophosphoryl
44	142.5	5.8	250 2 H83784	glycerophosphodiest
45	142	5.7	281 2 H86632	conserved hypothet

RESULT 2
 S55932 glycerophosphodiester phosphodiesterase (EC 3.1.4.46) - *Haemophilus influenzae* (strain 3
 N) [Alternative names: immunoglobulin D-binding protein; protein D
 C:Species: *Haemophilus influenzae*
 A:Variety: strain 3640
 C:Date: 15-feb-1996 #sequence_revision 01-Mar-1996 #text_change 08-Oct-1999
 C:Accession: S55932; S47334
 A:Reference number: S55931; MUID:7822043
 A:Accession: S55932
 A:Molecule type: DNA
 A:Residues: 1-364 <S0N>
 A:Cross references: EMBL:235657; NID:9525215; PIDN:CAA84716.1; PID:9525216
 A:Experimental source: strain 3640
 C:Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen

Query	Match	Score	Length	DB	2;
Best Local Matches	77.4% ; Score 1916 ; Pred. No. 2.7e-122 ; Mismatches 361 ; Conservative Matches 361 ;	99.2% ; 0 ; 0 ;	364 ; 0 ; 0 ;		
QY	1 MDPKTLAISLLAAGVLAGCSSHSSMMANTQMSDKIIIAHRSAGSYLPEHTLESKALAF	60			
Ddb	1 MKCUTLAISLLAAGVLAGCSSHSSMMANTQMSDKIIIAHRSAGSYLPEHTLESKALAF	60			
QY	61 QQADYLEQDLAMTKDGRLLVVIHDHFLDGLTLDVAKKPFHRKDGRRYYVDFTLKRIQSLE	120			
Ddb	61 QQADYLEQDLAMTKDGRLLVVIHDHFLDGLTLDVAKKPFHRKDGRRYYVDFTLKRIQSLE	120			
QY	121 MTENFETKDGKQAOQVNPRLWIKSHPRHIFTFEDIEFPIQCLESTGKKVGYPPIKAPW	180			
Ddb	121 MTENFETKDGKQAOQVNPRLWIKSHPRHIFTFEDIEFPIQCLESTGKKVGYPPIKAPW	180			
QY	181 PHHNGKQDIAETLKVLLKGYDICKTDMVYLOTFDNELKRIKTELLPONGMDLKIVQLI	240			
Ddb	181 FHHQNGKQDIAETLKVLLKGYDICKTDMVYLOTFDNELKRIKTELLPONGMDLKIVQLI	240			
QY	241 AYTDWKETQEKDPKGYWWNNNNDDMWFPGAMAEVWVYADGVPGKMYLNVKEESKPDNTV	300			
Ddb	241 AYTDWKETQEKDPKGYWWNNNNDDMWFPGAMAEVWVYADGVPGKMYLNVKEESKPDNTV	300			
QY	301 YTPVKEVLAQYINVEVHPYPTVTKDALLPEFTDVNQMDALLNKSGATGVFTDPPDTGVEFL	360			
Ddb	301 YTPVKEVLAQYINVEVHPYPTVTKDALLPEFTDVNQMDALLNKSGATGVFTDPPDTGVEFL	360			
QY	361 KGIK 364				
Ddb	361 KGIK 364				

Query Match	Score 1914;	DB 2;	Length 364;
Best Local Similarity	99 %;	Pred. No. 3.7e-122;	
Matches	361;	Missmatches	0;
Y			
1	MDPKTLASLIIAAGTLAGCSSHSMNANTOMKSDKIIIAHRSAGSLPEHTLESKALAF	60	
b	1 MKLKTIALSLLAAGTLAGCSSHSMNANTOMKSDKIIIAHRSAGSLPEHTLESKALAF	60	
Y			
61	QQADYLEQDLAMTKQGRLVYTHDFLDGLTDVAKCPHRKDGRRYYVDFTLKEQSLE	120	
b	61 QQADYLEQDLAMTKQGRLVYTHDFLDGLTDVAKCPHRKDGRRYYVDFTLKEQSLE	120	
Y			
121	MTNENPETDKQKAQYTPNRPPLWKSFRIRHPEDEIEFQGLEKSTGKKGIVYPEIKAPW	180	
b	121 MTNENPETDKQKAQYTPNRPPLWKSFRIRHPEDEIEFQGLEKSTGKKGIVYPEIKAPW	180	
Y			
181	FHHQNGKDIAAETLKVLKKGIVYDKTDMYLTQFDNFELRIKTELLPONGMDLKLVLQI	240	
b	181 FHHQNGKDIAAETLKVLKKGIVYDKTDMYLTQFDNFELRIKTELLPONGMDLKLVLQI	240	

RESULT 4									
Y	241	AYTDKETQKDPKGWWNTNYDNMKPGAMAEVTKYADGVPGPQHMLVNKEESTPDPNIV	300	5934	Lyceorophodiester phosphodiesterase (EC 3.1.4.46) precursor - Haemophilus				
Y	241	AYTDKETQKDPKGWWNTNYDNMKPGAMAEVTKYADGVPGPQHMLVNKEESTPDPNIV	300		:Species: Haemophilus influenzae				
Y	301	YPLVYKELAQVNVEHPTVTKDALPEFFTVDVNQNYDALINKSGATGVFTDPDGVEFL	360		:Variety: Strain Bagan; strain HK695				
Y	301	YPLVYKELAQVNVEHPTVTKDALPAFFTVDVNQNYDALINKSGATGVFTDPDGVEFL	360		:Date: 15-Feb-1996 #sequence revision 01-Mar-1996 #text_change 09-Jul-2004				
Y	361	KGIK	364		:Accession: S59934; Song, X.M.; Forbgren, A.; Janbon, H.				
Y	361	KGIK	364		:Title: The gene encoding protein D (hpd) is highly conserved among Haemophilus				
Y					:Reference number: S59931; PMID: 9512210; PMID: 1822043				
Y					:Molecule type: DNA				
Y					:Cross-references: UNIPROT:Q06282; EMBL:Z35659; PIDN:CAA84718.1;				
Y					:Experimental source: strain Bagan				
Y					:Accession: S59935				
Y					:Molecule type: DNA				
Y					:Residues: 1-364 <SOW>				
Y					:Cross-references: EMBL:Z35660; PID:9525221; PIDN:CAA84719.1; PID:9525222				
Y					:Experimental source: strain HK695				
Y					:Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen				
Y					:1-18/Domain: signal sequence #status predicted <SIG>				
Y					:19-364/Product: protein D #status predicted <MAT>				
Y					Query Match	77.1%	Score 1910;	DB 2;	Length 364;
Y					Best Local Similarity	98.9%	Pred. No. 6-9e-122;		
Y					Matches 360;	Conservative 0;	Msmatches 4;	Indels 0;	Gaps 0
Y	1	MDPKTLSSLAAGVLAGCSSHSSMANTONKSDDKLIIAHRGASGYLPEHTLESKALAF	60						
Y	1	MRLKTLSSLAAGVLAGCSSHSSMANTONKSDDKLIIAHRGASGYLPEHTLESKALAF	60						
Y	61	QADYLEQDLAMTKGCRIVVTHDHLGTDVAKKEPHRHKDGGYYVDFTLKETQSLE	120						
Y	61	QADYLEQDLAMTKGCRIVVTHDHLGTDVAKKEPHRHKDGGYYVDFTLKETQSLE	120						
Y	121	MTENPETDKGKQAAQYPNRPLWKSHFR1HTFEDTEFLQGLEKSTGKVGYIYPPEKAPW	180						

Db	121	MTENFETKDGKQAAQVYNRFPLWKSFRHPIHTEDEIEPQSLBKSTGKVKVGIYPEIKAPW	180	RESULT 6
Qy	181	FHHQNGKDIAAETLKVKYGYDKKCTDMVYIQTDFNBLKRKTELLPQMNDLKLVQLI	240	glycerophosphodiester phosphodiesterase (EC 3.1.4.46) - Haemophilus influenzae (strain N
N;Alternate names: immunoglobulin D-binding protein; protein D				
Db	181	FHHQNGKDIAAETLKVKYGYDKKCTDMVYIQTDFNBLKRKTELLPQMNDLKLVQLI	240	C;Species: Haemophilus influenzae
Qy	241	AYTDWKETQEKDPKGYVNNYDWMFKPGAMAEVVKYADGVPGWMTMLVNGESEKPDNIV	300	C;Variety: strain 6-7626
C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 08-Oct-1999				
C;Accession: S59933; S47335				
R;Song, X.M.; Forsgren, A.; Janson, H.				
A;Title: The gene encoding protein D (hpd) is highly conserved among Haemophilus influenzae				
A;Infect. Immun. 63, 696-699, 1995				
A;Reference number: S59931; MUID:95122210; PMID:7822043				
A;Accession: S59933				
A;Molecule type: DNA				
A;Residues: 1-364 <SON>				
A;Cross-references: EMBL:Z35658; NID:9525217; PID:CAA8471.1; PID:9525218				
A;Experimental source: strain 6-7626				
C;Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen				
Query Match	76.1%	Score 1885; DB 2; Length 364;		
Best Local Similarity	97.5%	Pred. No. 3-4e-120; Mismatches 4; Indels 5; Gaps 0;		
Matches 355; Conservative				
Qy	1	MDPKTIALSLLAAGVLAGCSHSSMANTONMSKDKIIIAHRSAGSYLPEHTLESKALAF	60	
Db	1	MRLKCTIALSLLAAGVLAGCSHSSMANTONMSKDKIIIAHRSAGSYLPEHTLESKALAF	60	
Qy	61	QADYLDQDLAMTKGRLVYTHDHLGLTDVAKCPHRKDGRRYVIDFTLKELOS	120	
Db	61	QADYLDQDLAMTKGRLVYTHDHLGLTDVAKCPHRKDGRRYVIDFTLKEQS	120	
Qy	121	MTENFETKDGKQAQVYPNRPFLLWQSHRINTFEDTEPIQLEKSTGKVKYIYPEIKAPW	180	
Db	121	MTENFETKDGKQAQVYPNRPFLLWQSHRINTFEDTEPIQLEKSTGKVKYIYPEIKAPW	180	
Qy	181	FHHQNGKDIABTLKVKYGYDKKTDMYLQTPDENELKRKTELLPQMGMDLKLVQLI	240	
Db	181	FHHQNGKDIABTLKVKYGYDKKTDMYLQTPDENELKRKTELLPQMGMDLKLVQLI	240	
Qy	241	AYTDWKETQEKDPKGYVNNYDWMFKPGAMAEVVKYADGVPGWMTMLVNGESEKPDNIV	300	
Db	241	AYTDWKETQEKDPKGYVNNYDWMFKPGAMAEVVKYADGVPGWMTMLVNGESEKPDNIV	300	
Qy	301	YTPLVKELAQTQVYHPTVTKDAPFTDYNQMDALLNKSGATGVFTDPDTGVEFL	360	
Db	301	YTPLVKELAQTQVYHPTVTKDAPFTDYNQMDALLNKSGATGVFTDPDTGVEFL	360	
Qy	361	KGIK 364		
Db	361	KGIE 364		
RESULT 5				
G64086				
glycerophosphodiester phosphodiesterase (EC 3.1.4.46) - Haemophilus influenzae (strain R				
N;Alternate names: IgD-binding protein; protein D				
C;Species: Haemophilus influenzae				
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004				
C;Accession: G61086				
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkinness, E.F.; Kerlavage, R.				
; Gocayne, J.D.; Scott, J.; Shirlow, R.; Liu, C.; Kelley, J.; Weidman, J.				
; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geohagen, N.S.M.				
; Science 269, 496-512, 1995				
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.				
A;Title: Whole genome random sequencing and assembly of Haemophilus influenzae Rd.				
A;Reference number: A64000; MUID:95150630; PMID:7542800				
A;Accession: G64086				
A;Status: nucleic acid sequence not shown; translation not shown				
A;Molecule type: DNA				
A;Cross-references: UNIPROT:Q06282; GB:U32751; GB:U42023; NID:93212200; PID: AAC22348.1/				
C;Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen				
Query Match	76.8%	Score 1902; DB 2; Length 364;		
Best Local Similarity	98.4%	Pred. No. 2; 4e-121; Mismatches 2; Indels 0; Gaps 0;		
Matches 358; Conservative				
Qy	1	MDPKTIALSLLAAGVLAGCSHSSMANTONMSKDKIIIAHRSAGSYLPEHTLESKALAF	60	
Db	1	MDKLTIALSLLAAGVLAGCSHSSMANTONMSKDKIIIAHRSAGSYLPEHTLESKALAF	60	
Qy	61	QADYLDQDLAMTKGRLVYTHDHLGLTDVAKCPHRKDGRRYVIDFTLKELOS	120	
Db	61	QHSDYLDQDLAMTKGRLVYTHDHLGLTDVAKCPHRKDGRRYVIDFTLKEQS	120	
Qy	121	MTENFETKDGKQAAQVYNRFPLWKSFRHPIHTEDEIEPQSLBKSTGKVKVGIYPEIKAPW	180	RESULT 7
Db	121	MTENFETKDGKQAAQVYNRFPLWKSFRHPIHTEDEIEPQSLBKSTGKVKVGIYPEIKAPW	180	glycerophosphodiester phosphodiesterase (EC 3.1.4.46) - Haemophilus influenzae (strain N
Qy	181	FHHQNGKDIAAETLKVKYGYDKKCTDMVYIQTDFNBLKRKTELLPQMNDLKLVQLI	240	N;Alternate names: IgD-binding protein; protein D
Db	181	FHHQNGKDIAAETLKVKYGYDKKCTDMVYIQTDFNBLKRKTELLPQMNDLKLVQLI	240	C;Species: Haemophilus influenzae
Qy	241	AYTDWKETQEKDPKGYVNNYDWMFKPGAMAEVVKYADGVPGWMTMLVNGESEKPDNIV	300	C;Variety: strain NTC 9468
Db	241	AYTDWKETQEKDPKGYVNNYDWMFKPGAMAEVVKYADGVPGWMTMLVNGESEKPDNIV	300	C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 08-Oct-1999
Qy	301	YTPLVKELAQTQVYHPTVTKDAPFTDYNQMDALLNKSGATGVFTDPDTGVEFL	360	C;Accession: S59936; S47338
Db	301	YTPLVKELAQTQVYHPTVTKDAPFTDYNQMDALLNKSGATGVFTDPDTGVEFL	360	R;Song, X.M.; Forsgren, A.; Janson, H.
Qy	361	KGIK 364		A;Title: The gene encoding protein D (hpd) is highly conserved among Haemophilus influenzae
Db	361	KGIE 364		A;Infect. Immun. 63, 696-699, 1995
Qy	361	KGIK 364		A;Reference number: S59931; MUID:95122210; PMID:7822043
Db	361	KGIE 364		A;Molecule type: DNA
A;Residues: 1-364 <SON>				A;Cross-references: EMBL:Z35651; NID:9525222; PID:CAA84720.1; PID:9525224
A;Experimental source: strain NTC 8468				C;Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen

Query Match	76.1%;	Score 1885;	DB 2;	Length 364;		Qy	244	DWKEHQEKDPKGWVNNYDWMFKPQGAMAEVVVKYADGVPGWYMLVNKEESKPDMIVYTP 303
Best Local Similarity	97.3%;	Pred. No. 3.4e-120;				Db	240	DWNEEQQKQDPGSWVNNYDWMFKPQGAMAEVVVKYADGVPGWYMLVNKEESKPDMIVYTP 303
Matches 355; Conservative 2;	Mismatches 0;	Indels 0;	Gaps 0;			Qy	304	LVKELAQYNNVBPVTPVKDALPEFTDYNQMDALLNSGATGVFTDPEPDGVEFL 360
Qy	1	MDPKTLASSLAAVGLAGCSSSNMANTOMSKDIIIAHKGASGLPPEHTLESKALATA 60				Db	299	MVQDQAQNKLVHVHPVRSKLPETTDVNQLDLYNKAGVNGLFTFDPKAVKPL 355
Db	1	MKLTKLSSLAAVGLAGCSSSNMANTOMSKHIIIAHKGASGLPPEHTLESKALATA 60						
Qy	61	QADYIEQDLAMTKDQGRLVWVHDHELDGTDVAKKEPCHRHKDGRYYVLFETLKERQSE 120						
Db	61	QADYIEQDLAMTKDQGRLVWVHDHELDGTDVAKKEPCHRHKDGRYYVLFETLKERQSE 120						
Qy	121	MTENFETKDGKQAOQYPNRPLWKSFRHTRPDELETFIQCLEKSTGKVGYIPRIKAPW 180						
Db	121	MTENFETKDGKQAOQYPNRPLWKSFRHTRPDELETFIQCLEKSTGKVGYIPRIKAPW 180						
Qy	181	FHHQNGKDIAAETKLVKKYGDCKTDMVYLOQTDFNEELKRKTELLPQMGMDLKLVQLI 240						
Db	181	FHHQNGKDIAAETKLVKKYGDCKTDMVYLOQTDFNEELKRKTELLPQMGMDLKLVQLI 240						
Qy	241	AYTDWKETQEDPKGWWVNNYDWMFKPQGAMAEVVVKYADGVPGWYMLVNKEESKPDMIV 300						
Db	241	AYTDWKETQEDPKGWWVNNYDWMFKPQGAMAEVVVKYADGVPGWYMLVNKEESKPDMIV 300						
Qy	301	YTPLVKELAQYNNVBPVTPVKDALPEFTDYNQMDALLNSGATGVFTDPEPDGVEFL 360						
Db	301	YTPLVKELAQYNNVBPVTPVKDALPEFTDYNQMDALLNSGATGVFTDPEPDGVEFL 360						
Qy	361	KGIK 364						
Db	361	KGIK 364						
Db	*							
RESULT 9								
D91019	9	glycerophosphodiester phosphodiesterase [imported] - Escherichia coli (strain O157:H7, <i>E. coli</i>)				Qy	4	KTLAISLLAAGVLAGCSSHSSNNMANTOMSKDIIIAHKGASGLPPEHTLESKALAFQA 63
C; Species: Escherichia coli		C; Accession: D91019				Db	6	KNLSMAMMSTIVMGSSAAAAD-----SNEKIVIAHGRASGVLPEHTLPKAKMAYAQA 59
C; Accession: D91019		C; Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004				Qy	64	DYLEQDLAMTKDGRLLVHDFLQGLDTDVKHFRPHRKDGYYVLFETLKEISLEMTE 123
R; Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Han, C.G.		Db	60	DYLEQDLVMTKDDHIVLHDHYLDRVTDAFRPFRARDKGRYYAIDFTLDEIISLKFT 119				
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.		Qy	124	NFETDKGKOAQYVNPREFLWKSFRHTRPDEFTQGLEKSTGKVGKYPEIKAPWFFH 183				
DNA Res. 8, 11-22, 2001		Db	120	GFDIENGKVKVQTPRFPMGSKDFRVTFFEEIEVQGLNHSIGNIGNIPEIKAPWFFH 179				
A; Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene		Qy	184	QNGKUAAETKLVKKYGDCKTDMVYLOQFDENELKRKTELLPQMGMDLKLVOLIAYT 243				
A; Reference number: A99629; PMID:21156231; PMID:11287936		Db	180	QEGKIDAAKTLLEVLYKYGKDDKRYLOCFDADELKRKNELEFKGMNDLNLVOLIAYT 239				
A; Accession: D91019		Qy	244	DWKEHQEKDPKGWWVNNYDWMFKPQGAMAEVVVKYADGVPGWYMLVNKEESKPDMIVYTP 303				
A; Status: preliminary		Db	240	DWNEEQQKQDPGSWVNNYDWMFKPQGAMAEVVVKYADGVPGWYMLVNKEESKPDMIVYTP 303				
A; Molecule type: DNA		Qy	304	LVKELAQYNNVBPVTPVKDALPEFTDYNQMDALLNSGATGVFTDPEPDGVEFL 360				
A; Residues: 1-358 <TOM>		Db	299	MVQDQAQNKLVHVHPVRSKLPETTDVNQLDLYNKAGVNGLFTFDPKAVKPL 355				
A; Cross-references: UNIPROT:Q8XE17; GB:BA000007; PID:BAB36547_1; PID:913362594; GSPDB:G								
A; Experimental source: strain O157:H7, <i>Escherichia coli</i> O157:H7 and gene								
A; Reference number: A99629; PMID:21156231; PMID:11287936								
A; Accession: D91019								
A; Status: preliminary								
A; Molecule type: DNA								
A; Residues: 1-358 <TOM>								
A; Cross-references: UNIPROT:Q8XE17; GB:BA000007; PID:BAB36547_1; PID:913362594; GSPDB:G								
A; Experimental source: strain O157:H7, <i>Escherichia coli</i> O157:H7 and gene								
A; Reference number: A99629; PMID:21156231; PMID:11287936								
A; Accession: D91019								
RESULT 10								
D91019	10	glycerophosphodiester phosphodiesterase (EC 3.1.4.46) precursor, periplasmic - Escherich				Qy	4	KTLAISLLAAGVLAGCSSHSSNNMANTOMSKDIIIAHKGASGLPPEHTLESKALAFQA 63
C; Species: Escherichia coli		C; Accession: S15945				Db	6	KNLSMAMMSTIVMGSSAAAAD-----SNEKIVIAHGRASGVLPEHTLPKAKMAYAQA 59
C; Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 09-Jul-2004		R; Tommasen, J.; Englmaier, K.; Cole, S.T.; Overduin, P.; Larson, T.J.; Boos, W.				Qy	64	DYLEQDLAMTKDGRLLVHDFLQGLDTDVKHFRPHRKDGYYVLFETLKEISLEMTE 123
R; Tommasen, J.; Englmaier, K.; Cole, S.T.; Overduin, P.; Larson, T.J.; Boos, W.		Db	60	DYLEQDLVMTKDDHIVLHDHYLDRVTDAFRPFRARDKGRYYAIDFTLDEIISLKFT 119				
Mol. Gen. Genet. 246, 321-327, 1991		Qy	124	NFETDKGKOAQYVNPREFLWKSFRHTRPDEFTQGLEKSTGKVGKYPEIKAPWFFH 183				
A; Title: Characterization of two genes, <i>g1pQ</i> and <i>u1pQ</i> , encoding glycerophosphoryl dieste		Db	120	GFDIENGKVKVQTPRFPMGSKDFRVTFFEEIEVQGLNHSIGNIGNIPEIKAPWFFH 179				
A; Reference number: S15945; PMID:91238812; PMID:1851953		Qy	184	QNGKUAAETKLVKKYGDCKTDMVYLOQFDENELKRKTELLPQMGMDLKLVOLIAYT 243				
A; Molecule type: DNA		Db	180	QEGKIDAAKTLLEVLYKYGKDDKRYLOCFDADELKRKNELEFKGMNDLNLVOLIAYT 239				
A; Residues: 1-358 <TOM>		Qy	244	DWKEHQEKDPKGWWVNNYDWMFKPQGAMAEVVVKYADGVPGWYMLVNKEESKPDMIVYTP 303				
A; Cross-references: UNIPROT:D09394; EMBL:X56907; PID:941580; PMID:CAA40223_1; PID:941581		Db	240	DWNEEQQKQDPGSWVNNYDWMFKPQGAMAEVVVKYADGVPGWYMLVNKEESKPDMIVYTP 303				
R; Englmaier, K.; Boos, W.; Cole, S.T.								

A;Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AH0791
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-356 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD07514.1; PID:g16503507; GSPPDB:GN00176
C;Genetics:
A;Gene: Srv2511
C;Keywords: phosphoritic diester hydrolase

Qy	121	MTEENFE-TKDGKQAOQVYPNRNPFPLWKSHFRHTDEIEFQGLEKSTGKVGIVYPPBIKAP	179	glycerophosphoryl diester phosphodiesterase VCA0136 [Imported] - <i>Vibrio cholerae</i> (strain
Db	125	FTWGPFDIDKNGKQVQSPNRPMPGKDFKRTFQVETIEFQGLNSTGNTGIVYPPBIKAP	184	Species: <i>Vibrio cholerae</i>
Qy	180	WFFHONGKQDIAETLKVKKYGYDKEETDMYLQTFDNEELKRKITEELLPGMGLKLVOL	239	C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession:	H82497	J.-F. Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Heidelberg, J.; Nelson, W.C.; Qin, H.; Dragoi, I.; Sellers, P.; Richardson, D.; Brinkac, M.; Vamathevan, J.; Fraser, C.M.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.		
Db	185	WFFHQBESKDFDSTKVLTEKEYTNGKDFVTLQSFVNELKRKNELEPGMGLKLVOL	244	Nature 406, 477-483, 2000
Qy	240	IAYTDWKETQEDPKGKWWVNNYDWMFKPGAMAEVVKYADGVPGKWMVNLVKEESKPDNI	299	A;Title: DNA sequence of both chromosomes of the cholera pathogen <i>Vibrio cholerae</i> .
Db	245	IAYTDWNEETYEOLDPDKGKWWVNNYDWMFKPGAMKEYAHADGIGDQHMLVVKTSTAKD-I	303	A;Reference number: A82035; MUID: 20406833; PMID: 10952301
Qy	300	VTPVYKELAQNVEVHPVTPYTKDALPEFTDVNQMDALLNKSDATGWTGPTDPPDTGVBF	359	A;Status: preliminary
Db	304	KLTDLYKEAHASNSMHTHPTTRADRPKYATDINGQFLLYNQAOVQDGVFTDPPQIQCFT	363	A;Cross-references: UNIPROT:Q9KRN30; GB:AE004354; GB:AE003853; NID:99657512; PIDN:AAF9604
Qy	360	LK 361		A;Experimental source: serogroup O1; strain N16561; biotype El Tor
Db	364	LQ 365		C;Genetics:
Qy				A;Gene: VCA0136
Db				A;Map position: 2
RESULT 13				
P71346				
probable glycerophosphoryl diester phosphodiesterase (gpl0) - <i>syphilis</i> spirochete				
C;Species:	<i>Treponema pallidum</i> subsp. <i>pallidum</i> (syphilis spirochete)			
C;Accession:	P71346			
C;Date:	24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004			
R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwinn, J.; Weilak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDaniel, J.; Weidman, J.; Smith, H.O.; Venter, J.C.				
Science 281, 375-388, 1998				
A;Title: Complete genome sequence of <i>Treponema pallidum</i> , the <i>syphilis</i> spirochete.				
A;Reference number: A71250; MUID: 93332770; PMID: 9665876				
A;Accession: P71346				
A;Status: preliminary; nucleic acid sequence not shown; translation not shown				
A;Molecule type: DNA				
A;Residues: 1-156 <CDS>				
A;Cross-references: UNIPROT:O30405; GB:AE001207; GB:AB000520; MUID: 9322526; PIDN: AAC6524				
A;Experimental source: strain Nichols				
C;Genetics:				
A;Gene: TP0257				
RESULT 14				
P71346				
probable glycerophosphoryl diester phosphodiesterase (gpl0) - <i>Nostoc</i> sp. (strain PCC 7120)				
C;Species:	<i>Nostoc</i> sp. PCC 7120			
A;Note: <i>Nostoc</i> sp. strain PCC 7120 is a synonym of <i>Anabaena</i> sp. strain PCC 7120				
C;Accession:	AC1841			
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchi, S.; Nakazaki, N.; Shimpoo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S.				
DNA Res. 8, 205-213, 2001				
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium <i>Anabaena</i>				
A;Accession: AC1841				
A;Status: preliminary				
A;Molecule type: DNA				
A;Residues: 1-1027 <CDS>				
A;Cross-references: UNIPROT:Q8Z028; GB:BA000019; PIDN: BAB7799-1; PMID: 11759840				
A;Experimental source: strain PCC 7120				
A;Gene: all0275				
RESULT 15				
H82497				
glycerophosphoryl diester phosphodiesterase [Imported] - <i>Nostoc</i> sp. (strain PCC 7120)				
C;Species:	Nostoc sp. PCC 7120			
C;Accession:	AC1841			
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchi, S.; Nakazaki, N.; Shimpoo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S.				
DNA Res. 8, 205-213, 2001				
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium <i>Anabaena</i>				
A;Accession: AC1841				
A;Status: preliminary				
A;Molecule type: DNA				
A;Residues: 1-1027 <CDS>				
A;Cross-references: UNIPROT:Q8Z028; GB:BA000019; PIDN: BAB7799-1; PMID: 11759840				
A;Experimental source: strain PCC 7120				
A;Gene: all0275				
RESULT 16				
H82497				
glycerophosphoryl diester phosphodiesterase [Imported] - <i>Nostoc</i> sp. (strain PCC 7120)				
C;Species:	Nostoc sp. PCC 7120			
C;Accession:	AC1841			
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchi, S.; Nakazaki, N.; Shimpoo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S.				
DNA Res. 8, 205-213, 2001				
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium <i>Anabaena</i>				
A;Accession: AC1841				
A;Status: preliminary				
A;Molecule type: DNA				
A;Residues: 1-1027 <CDS>				
A;Cross-references: UNIPROT:Q8Z028; GB:BA000019; PIDN: BAB7799-1; PMID: 11759840				
A;Experimental source: strain PCC 7120				
A;Gene: all0275				

Query Match Score 425; DB 2; Length 1027;
 Best Local Similarity 29.4%; Pred. No. 9.6e-2;
 Matches 137; Conservative 66; Mismatches 147; Indels 116; Gaps 20;

Qy 37 IIAHARGASGYLPBPHTLSKALLAFAQQADYLBODLAMTKGRLUVIHD-----H 84
 Db 180 VIGHRGASGFRFENTLESYKLAIEQGADPFLATYKDGVLJARHEPALAVNADGSVN 239

Qy 85 FUDGLTDYAK--KPFPHRIRK--DGR----YVVIDFTLKEOTSLENTEFETKGQAOV 135
 Db 240 FSNTTTINYQIAKFSRKLTKVNLDGTETGWAEDFTLAETKELRAIERTL----- 289

Qy 136 YPNRFPLWKSFRHFRHTPDEIEPIQGLEKSTGKVGTYPEIKAQWPFTHQ----- 184
 Db 290 -PFRDOSNTNGOFTIPTLAELIIDLVKQVBAETGKKGIGIYPETKGHTYFAQEATYVGTTTEKI 348

Qy 185 --NGKDIAAETLKVLUKGYDKKTD--MVYLQTFDFNELKRIKTELLPQMGMIDLKUVQLI 240
 Db 349 NRNISQIILIDTLKA----NNTFTDPSRISIQTOSFEVGNLKEHDTIMPAGVDIPLVQLF 402

Qy 241 AYTDWKEETQEKDDEKGYWWNN-YDMMFK-----PGAMAEVVKYADGUGPQWML 288
 Db 403 -----DAVIDDINGRLELTERPFDIYSGDTPTYGDLRTPAGIAETAYAIGIGPKWRI 456

Qy 289 VNKEESKPDN-----IVYTPYKELAQNYNEVHPYPTVRKDALPFF 328
 Db 457 VSVRGTDANNQGADDYNGDGAVNDADKTLIPLPTTLVQDAHNVGLOVHPYTFRDE---ER 513

Qy 329 FTDVNQW-----YDALLINKSATGVTDFPPTGVEFLKGKSMDGGRAGVALVRSDYKL 382
 Db 514 YLAANYQQNPELEYQQLF-QLGVDALFTDFPIT-ADRVRDRLSLPGNN---IVRSP--- 564

Qy 383 YNKNSSSNISTURKNGLGEHRARAMDGGKAGVALVRSDFKFYEDANGT 428
 Db 565 QNPDVLSGDAFANLG---GSRGEGG---AINASKFKLYMLLEG 603

Search completed: June 22, 2005, 11:41:29
 Job time: 43 secs

THIS PAGE BLANK (USPTO)

Scoring table:	BLOSUM62	ALIGMENTS				
Gapop:	10.0 , Gapext 0.5					
Searched:	1612378 seqs, 512079187 residues					
Total number of hits satisfying chosen parameters:	1612378					
Minimum DB seq length:	0					
Maximum DB seq length:	2000000000					
Post-processing:	Minimum Match 0% Maximum Match 100%					
	Listing first 45 summaries					
Database :	UniProt_03-* 1: uniprot_sprot: 2: uniprot_trembl: *					
Pred.	No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.					
SUMMARIES						
Result No.	Score	Query	Match	Length	DB ID	Description
1	1902	76.8	364	1	GLPQ_HAEIN	Q06282 haemophilus
2	1577	63.7	358	2	Q9CL04	Q9CL04 pastoreilla
3	1501.5	60.6	368	2	Q65R11	Q65R11 manneheimia
4	1498.5	60.5	357	2	Q8RHT6	Q8RHT6 fusobacteri
5	13.63	55.0	309	2	Q7P6W6	Q7P6W6 fusobacteri
6	1261.5	50.9	358	2	Q8XE17	Q8XE17 escherichia
7	1251.5	50.6	358	2	Q8CTV6	Q8CTV6 escherichia
8	1251.5	50.5	358	1	GLPQ_ECOLI	P09394 escherichia
9	1551.5	50.5	358	2	Q7UC72	Q7UC72 shigella fl
10	1231.5	50.5	377	2	Q83QU2	Q83QU2 shigella fl
11	1248.0	50.4	361	2	Q6CZI4	Q6CZI4 erwinia car
12	1235	49.9	356	2	Q82NG7	Q82NG7 salmonella
13	1233	49.8	356	2	Q82556	Q82556 salmonella
14	1210.5	48.9	358	2	Q7MZV7	Q7MZV7 photorhabdu
15	1210	48.8	371	2	Q66FX7	Q66FX7 yersinia pe
16	1210	48.8	371	2	Q8ZAH4	Q8ZAH4 yersinia pe
17	1210	48.8	372	2	Q8D1J5	Q8D1J5 yersinia pe
18	1082	43.7	356	2	Q94X5U6	Q94X5U6 treponema p
19	1074	43.4	356	1	GLPQ_TREPBA	P030405 treponema p
20	1074	43.4	356	2	Q71U57	Q71U57 treponema s
21	1074	43.4	356	2	Q71U58	Q71U58 treponema p
22	1074	43.4	356	2	Q71U59	Q71U59 treponema p
23	1074	43.4	356	2	Q71U64	Q71U64 treponema p
24	1072.5	43.3	341	2	Q7QMY5	Q7QMY5 anophelles g
25	1026	41.4	360	2	Q61JT65	Q61JT65 photobacter
26	9875	39.9	351	2	Q87M77	Q87M77 vibrio para
27	979	39.5	352	2	Q7NDR9	Q7NDR9 vibrio vuln
28	979	39.5	352	2	Q8D6W7	Q8D6W7 vibrio vuln
29	957	38.6	371	2	Q9kN30	Q9kN30 vibrio chol
30	839.5	33.9	338	2	Q9F8D4	Q9F8D4 borrelia tu
31	836	33.8	337	2	Q9F8D3	Q9F8D3 borrelia co

RA	Song X.-M., Forssgren A., Jansson H.,	Forssgren A., Jansson H.,	191	191	T -> A (in strain Bagan, strain 3639, strain 6- strain 3640, strain NCTC 8468, strain 6- strain HK695 and strain Minna).
RT	Haemophilus influenzae type b and nontypeable strains. ";				P -> S (in strain 6-7626).
RL	Infect. Immun. 63:696-699(1995).				Q -> K (in strain 6-7626).
RN	[5]				B -> A (in strain Bagan, strain 3639, strain NCTC 8468, strain 6-7626, strain HK695 and strain Minna).
RP	CHARACTERIZATION.				A -> V (in strain Bagan, strain 3640, strain HK695 and strain Minna).
RC	STRAIN=NTHI_772;				K -> E (in strain 6-7626).
RX	MEIDLINE=92192801; PubMed=1548059;				K -> E (in strain 6-7626).
RA	Jansson H., Heden L.-O., Forssgren A.;				A6079B3ABF70B820 CRC64;
RT	"Protein D, the immunoglobulin D-binding protein of Haemophilus influenzae, is a lipoprotein.";				
RL	Infect. Immun. 60:1336-1342(1992).				
-I -	FUNCTION: Glycerophosphoryl diester phosphodiesterase hydrolyzes deacylated phospholipids to G3P and the corresponding alcohols.				
CC	Has a specific affinity for human immunoglobulin D myeloma protein.				
CC	-I - CATALYTIC ACTIVITY: A glycerophosphodiester + H(2)O = an alcohol + sn-glycerol 3-phosphate.				
CC	-I - SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid anchor.				
CC	-I - PTM: Contains both ester- and amide-linked fatty acids.				
CC	-I - MISCELLANEOUS: The sequence shown is that of strains NTHI 772 and RD / KW20.				
CC	-I - SIMILARITY: Belongs to the glycerophosphoryl diester phosphodiesterase family.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL: M332751; AAC22348.1; -;				Q9CJ04 PRELIMINARY; PRT; 358 AA.
DR	EMBL: M337487; AAA24998.1; -;				Q9CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	EMBL: L12445; AAA24999.1; -;				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	• PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
KW	Complete proteome; Glycerol metabolism; Hydrolase; Lipoprotein; Outer membrane; Palmitate; Signal.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	SIGNAL 1 18				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	CHAIN 19 364				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	LIPID 19 19				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	LIPID 19 19				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 13 13				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 16 16				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 25 25				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 28 28				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 34 34				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 62 62				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	PIR; G64086; G64086.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	TIGR; H10089; -;				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	IntPro; IPR004129; GDPD.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	PFam; PF01009; GDPD. 1.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
KW	Complete proteome; Glycerol metabolism; Hydrolase; Lipoprotein; Outer membrane; Palmitate; Signal.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	SIGNAL 1 18				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	CHAIN 19 364				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	LIPID 19 19				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	LIPID 19 19				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 13 13				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 16 16				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 25 25				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 28 28				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 34 34				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 62 62				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	PIR; G64086; G64086.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	TIGR; H10089; -;				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	IntPro; IPR004129; GDPD.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	PFam; PF01009; GDPD. 1.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
KW	Complete proteome; Glycerol metabolism; Hydrolase; Lipoprotein; Outer membrane; Palmitate; Signal.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	SIGNAL 1 18				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	CHAIN 19 364				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	LIPID 19 19				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	LIPID 19 19				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 13 13				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 16 16				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 25 25				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 28 28				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 34 34				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 62 62				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	PIR; G64086; G64086.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	TIGR; H10089; -;				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	IntPro; IPR004129; GDPD.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	PFam; PF01009; GDPD. 1.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
KW	Complete proteome; Glycerol metabolism; Hydrolase; Lipoprotein; Outer membrane; Palmitate; Signal.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	SIGNAL 1 18				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	CHAIN 19 364				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	LIPID 19 19				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	LIPID 19 19				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 13 13				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 16 16				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 25 25				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 28 28				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 34 34				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 62 62				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	PIR; G64086; G64086.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	TIGR; H10089; -;				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	IntPro; IPR004129; GDPD.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	PFam; PF01009; GDPD. 1.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
KW	Complete proteome; Glycerol metabolism; Hydrolase; Lipoprotein; Outer membrane; Palmitate; Signal.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	SIGNAL 1 18				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	CHAIN 19 364				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	LIPID 19 19				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	LIPID 19 19				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 13 13				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 16 16				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 25 25				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 28 28				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 34 34				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 62 62				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	PIR; G64086; G64086.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	TIGR; H10089; -;				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	IntPro; IPR004129; GDPD.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	PFam; PF01009; GDPD. 1.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
KW	Complete proteome; Glycerol metabolism; Hydrolase; Lipoprotein; Outer membrane; Palmitate; Signal.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	SIGNAL 1 18				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	CHAIN 19 364				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	LIPID 19 19				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	LIPID 19 19				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 13 13				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 16 16				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 25 25				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 28 28				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 34 34				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 62 62				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	PIR; G64086; G64086.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	TIGR; H10089; -;				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	IntPro; IPR004129; GDPD.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	PFam; PF01009; GDPD. 1.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
KW	Complete proteome; Glycerol metabolism; Hydrolase; Lipoprotein; Outer membrane; Palmitate; Signal.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	SIGNAL 1 18				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	CHAIN 19 364				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	LIPID 19 19				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	LIPID 19 19				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 13 13				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 16 16				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 25 25				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 28 28				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 34 34				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 62 62				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	PIR; G64086; G64086.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	TIGR; H10089; -;				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	IntPro; IPR004129; GDPD.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	PFam; PF01009; GDPD. 1.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
KW	Complete proteome; Glycerol metabolism; Hydrolase; Lipoprotein; Outer membrane; Palmitate; Signal.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	SIGNAL 1 18				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	CHAIN 19 364				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	LIPID 19 19				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	LIPID 19 19				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 13 13				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 16 16				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 25 25				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 28 28				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 34 34				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	VARIANT 62 62				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
FT	PIR; G64086; G64086.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	TIGR; H10089; -;				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	IntPro; IPR004129; GDPD.				AC 09CJ04; ID 09CJ04; PRELIMINARY; PRT; 358 AA.
DR	PFam; PF01009; GDPD. 1.				

DR	EMBL: AB006181; AAK03528.1;	Qy	64 DYLEQDLAMTKDGRLLVYIHDHFLDGLTDVAKCKPFRHRKGRRYVLDFTLKEIOSLEMTE 123
DR	GO: 0006889; F: glycerophosphodiester phosphodiesterase act. . .; IEA.	Db	65 DYLEQDLAMTKDGRLLVYIHDHFLDGLTDVAKCKPFRHRKGRRYVLDFTLKEIOSLEMTE 124
DR	InterPro: IPR004129; GPD.	Qy	124 NFBTQDGKQAOVYPPRFPLWKSHPFRHTFEDIEFQGLEKSTGKVGTYPPBKAPMPFH 183
DR	PFam: PF03009; GPD; 1.	Db	125 NPKTENGKQVQVYPPRFPMWKSHPFTHTFEEBLEFQGLEKSTGKVGTYPPBKAPMPFH 184
KW	Complete proteome.		
SQ	SEQUENCE 358 AA; 41162 MW; 9BEF514318C30413 CRC64;		
Query Match	Score 1577; DB 2; Length 358;	Qy	184 QNGKDTAAETTKVLLKCYGYDKTDMVYLOQTDNEIURKIKTELLPONGMDKJLKVQIAYT 243
Best Local Similarity 82.3%; Pred. No. 2.2e-101; Matches 297; Conservative 27; Mismatches 31; Indels 6; Gaps 3;		Db	185 QEGKDAVATLKVQKGYTAKTKDTPYLTQTPDNENIURKIKDLPQMGMDVJLKVQIAYT 244
Db	1 MDPKTLAISLLAAGVLAGCSSHSSNNMANTQMSDKIIIAHRCASGVLPEHTLESKALAF 60	Qy	244 DWKETQKDQPGKWWNNYDMMKPGAMAEEVVKYADVGPGWMLVNEESKPDNIVYTP 303
Db	1 MCLKTL-VAIALSTLPTACSS-SMMRNN---DDKLIITIAHRCASGVLPEHTLESKALAF 54	Db	245 DWHETTEKNAQGKWWNYDMMKPGAMAEEVVKYADVGPGWMLVNEESKPDNIVYTP 304
Qy	61 QQADYLLQDQDLAMTKDGRLLVYIHDHFLDGLTDVAKCKPFRHRKGRTYVLDFTLKEIQLS 120	Qy	304 LVTKELQYNTVEVHPYPTVTKDPEFTDVMQMYDALLNKSGATGVTDPDTGQVFL 360
Db	55 QQADYLLQDQDLAMTKDGRLLVYIHDHFLDGLTDVAKCKPFRNRANKDGRYYVVDFTLKEIQLS 114	Db	305 MVADIAKTKHNLHPYTVRKDPEFTDVMQMYDALNHAATGJLPTDFPLAVKPL 361
Qy	121 MTENFETPKDGRKQAOVYNNRFPFLWKSHPFRHTFEDIEFQGLEKSTGKVGTYPEIKA 180		RESULT 4
Db	115 MTENFKVEYNGKQVQYTPNRFPLWKSHPFRHTFEDIEFQGLEKSTGKVGTYPEIKA 174	Q8RHT6	Q8RHT6 PRELIMINARY; PRT; 357 AA.
Qy	181 FHHONGKDIAAETTKVLLKCYGYDKTDMVYLOQTDNEIURKIKTELLPOMGNDIKLKVQLI 240	AC	Q8RHT6 PRELIMINARY; PRT; 357 AA.
Db	175 LHHKEGKDIAAETTKVLLKCYGYDTKODARVYLOQTDNEIURKIKTELLPOMGNDIKLKVQL 234	DT	01-JUN-2002 (T-EMBL;rel. 21, Created)
Qy	241 AYDWKETQKDQPGKWWNNYDMMKPGAMAEEVVKYADVGPGWMLVNEESKPDNIV 300	DT	01-JUN-2003 (T-EMBL;rel. 21, Last sequence update)
Db	235 AYDWHETTEKNAQGKWWNYDMMKPGAMAEEVVKYADVGPGWMLVNEESKPDNIV 294	DB	01-JUN-2003 (T-EMBL;rel. 24, Last annotation update)
Qy	301 YTPLVKELAQYNTVEVHPYPTVTKDPEFTDVMQMYDALLNKSGATGVTDPDTGQVFL 360	GN	OrderedLocusNames=FN1908;
Db	295 YTPLVKELAQYNTVEVHPYPTVRKDPEFTDVMQMYDALLNKSGATGVTDPDTGQVFL 354	OS	Fusobacterium nucleatum (subsp. nucleatum) .
Qy	361 K 361	OC	Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
Db	355 K 355	OC	Fusobacterium
RN	[1]	NCBI_TaxID=76356;	
RP	SEQUENCE FROM N.A.	NCBI_TaxID=76356;	
RC	STRAIN=ATCC 25586;	NCBI_TaxID=76356;	
RD	MEDLINE=2186334; PubMed=11889109;	NCBI_TaxID=76356;	
RX	DOI=10.1128/JB.184.7.2005-2018.2002;	NCBI_TaxID=76356;	
RA	Kapatral V, Anderson I, Ivanova N, Reznik G, Los T, Lykidis A., Bhattacharya A., Bartman A., Gardner W., Greshkin G., Zhu L., Vasileva O., Chu L., Kogut Y., Chaga O., Goltsman E., Berna A., Larsen N, D'Souza M., Walunas T., Pusch G., Haselkorn R., Fonstein M., Kyrpides N.C., Overbeek R.,	NCBI_TaxID=76356;	
RA	RT "Genome sequence and analysis of the oral bacterium Fusobacterium nucleatum strain ATCC 25586",	NCBI_TaxID=76356;	
RL	DOI=10.1128/JB.184.7.2005-2018.2002;	NCBI_TaxID=76356;	
J. Bacteriol. 184:2005-2018 (2002)			
EMBL; AB010493; AAEL94071;			
DR	DR GO; GO:0008889; F:glycerophosphodiester phosphodiesterase act. . .; IEA.		
DR	DR GO; GO:0006071; P:glycerol metabolism; IEA.		
DR	DR InterPro; IPR004129; GDPD.		
DR	DR Pfam; PF03009; GDPD; 1.		
KW	Complete proteome.		
SQ	SEQUENCE 357 AA; 41111 MW; BBECE3F3D3FD53B0 CRC64;		
Query Match	Score 60.5%; Score 1498.5%; DB 2; Length 357;		
Best Local Similarity 77.7%; Pred. No. 6.2e-96; Mismatches 35; Indels 3; Gaps 1;			
Matches 278; Conservative 35; Mismatches 42; Indels 3; Gaps 1;			
Qy	8 LSLLAAGVLAGCSSHSSNNMANTQMSDKIIIAHRCASGVLPEHTLESKALAFQQA 67	Db	68 QDLAQTQDGKWWYIHDHFLDGLTDVAKCKPFRHRKGRRYVLDFTLKEIOSLEMTE 127
RT	1 MKLKSCLVGLGILSSTALFA--ANGKIIIAHRCASGVLPEHTLESKALAFQQA 57	Db	58 QDLANSKDKLIVIDHFLGQTDVAKCKPFRHRKGRRYVLDFTWSELOPLEMNTFT 117
RT	RT "The genome sequence of the capnophilic rumen bacterium Mannheimia succiniciproducens",	Qy	128 KDGKQAOQVYPPRFPMWKSHPFRHTFEDIEFQGLEKSTGKVGTYPEIKA 187
RL	RT "The genome sequence of the capnophilic rumen bacterium Mannheimia succiniciproducens MBEL55B",	Db	118 KDGKQTAQVYPPRFPMWKSHPFRHTFEDIEFQGLEKSTGKVGTYPEIKA 177
DR	DR EM016827; AAU8599.1;	Qy	188 DIAETLKVLLKCYGDKKTDVMVYLOQTDENELKRKTELLPQMGMQLKLVQIAYTDWKB 147
SQ	SEQUENCE 368 AA; 42139 MW; 9F8A0D60B029873C CRC64;	Db	
Query Match	Score 60.6%; Score 1501.5%; DB 2; Length 368;		
Best Local Similarity 77.0%; Pred. No. 4e-96; Mismatches 44; Indels 1; Gaps 1;			
Matches 275; Conservative 37; Mismatches 44; Indels 1; Gaps 1;			
Qy	4 KTLALSLAAGVLAGCSSHSSNNMANTQMSDKIIIAHRCASGVLPEHTLESKALAFQQA 63	Db	
Db	6 KTLAIG-LAFLAIAACSSQTAQQTTPMNNQEKLVIAHRCASGVLPEHTLESKALAFQQA 64	Qy	

Db	299	MVQDAQQNKLVYHPYTVRSKLPFVTTDNQLYDALYNKAGVNGLETFDPDKAVKPL	355
RESULT 7			
QBCVV6	PRELIMINARY;	PRT;	358 AA.
ID	QBCVV6;		
AC			
DT	01-MAR-2003	(TREMBLref. 23, Created)	
DT	01-MAR-2003	(TREMBLref. 23, Last sequence update)	
DT	01-MAR-2003	(TREMBLref. 24, Last annotation update)	
DE		Glycerophosphoryl diester phosphodiesterase, periplasmic (EC 3.1.4.46).	
DB			
GN		Name=gipQ; OrderedLocusNames=c2780;	
OS		Escherichia coli O6.	
OC		Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;	
OC		Enterobacteriaceae; Escherichia.	
NCBI_TaxID	217992;		
RN		SEQUENCE FROM N.A.	
RP		STRAIN=K12 / CFT073 / ATCC 700928;	
RC		MDLINE=2238823;	
RX		PubMed=12471157; DOI=10.1073/pnas.252529799;	
RA		Welch R.A., Burland V., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Colrado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Boos W., Tommassen J., Biglmeier K., Cole S.T., Overduin P., Larson T.J., Boos W.;	
RA		Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;	
RA		"Characterization of two genes, gipQ and gipQ, encoding glycerophosphoryl diester phosphodiesterases of Escherichia coli K-12.";	
RA		Mobiley H.L.T., Donnenberg M.S., Blattner F.R.;	
RT		"Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli.";	
RT		Interro; IPR004129; GPD.	
RL		Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).	
DR		EMBL; AE016763; ANB1234.1; -	
DR		GO; GO:000889; F:glycerophosphodiester phosphodiesterase activity; IEA.	
DR		GO; GO:0016787; F:hydrolase activity; IEA.	
DR	1	DR; GO:0006071; P:glycerol metabolism; IEA.	
DR		Interro; IPR004129; GPD.	
DR		Pfam; PF03009; GPD; 1.	
KW		Complete proteome; Hydrolase.	
SQ		SEQUENCE 358 AA; 40826 MW; C633D35EA86591F8 CRC64;	
Qy		Query Match	
Matches	230;	Best Local Similarity	50.6%;
Qy		Conservative	Score 1254.5;
Qy		55;	DB 2;
Qy		Mismatches	Length 358;
Db		65;	Indels 2;
Db		2;	Gaps 2;
Qy	4	KTLAISLLAAAGVLAGCSCSHSNMANTOMSKDKIIIAHRGAGYLPLHTEBSKALAPQA	63
Db	6	KNLSSMAIMSTMIVGMSAAAD-----SNEKVIIAHRGAGYLPLHTEPLAKAMAYAQGA	59
Qy	64	DYLEQDIAAMTKDGRLVVYHDFLQDGLTVDVAKFPHHRKDRYVYDFTLKEIOSLEME	123
Db	60	DYLEQDIAAMTKDHLVHDHLDRYDPAKDRYDARQDRYDADTFLDBIKSUKTFE	119
Qy	124	NFETKDGKQAQYQVNPNEPFLWPKSHRHTFEDDEIEPQLEKSTGKVGIVPEIKAPWFH	183
Db	120	GFDIENCKVQYQVGPGRPMGKSDFRVHTFEEBIEFVQGLNHSITGKNGIYPEIKAPWFH	179
Qy	184	QNGKDAIAETLKLKKGYDQKTDMTLQTFDNEKLKRTPEMGLKLVQIAYT	243
Db	180	QEGKDAIAAKTLEVLKCYGTYGDKYQLOCFDADEKLKRNVLPEKGMQLNVLQIAYT	239
Qy	244	DWKETQKDKPQKSYQWNTYNDMFKPGAMAEMVYKVAUGVGPQWYLYNKESERKPDNTVYTP	303
Db	240	DWNETQKQDPSWVNTYNDMFKPGAMQVAEYQGDPYHML-EETSGPGNKLKG	298
Qy	304	LYKELAQYVNVBHPYTVRKAPEFFTDVNQMDYALLNKSATGTVYDFTDPDGKFL	360
Db	299	MVQDAQQNKLVYHPYTVRSKLPFVTTDNQLYDALYNKAGVNGLETFDPDKAVKPL	355
RESULT 8			
GLPQ_ECOLI	STANDARD;	PRT;	358 AA.
ID	GLPQ_ECOLI		
AC			
DT	01-MAR-1989	(Rel. 10, Created)	
DT	01-NOV-1991	(Rel. 20, Last sequence update)	
DT	25-JAN-2005	(Rel. 46, Last annotation update)	

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch.

Qy	244 DWKETQKPKGTYWNNYDMMFKPGAMAEVVKYADGVGPQWMLVNKEESKPDNIVTP 303	Qy	184 QNGKDIAAETLKVKKYGYDKKCTDMVYLOTFDNELKRKTTELPQMGMDLKLVOLIAYT 243
Db	237 DNNETQKQPKGRWVNNYDMMFKQAMKVAEYDGIGDYYHLY-AE8STKRNKLTG 295	Db	177 QEGKDIAAETLKVKKYGYDKKCTDMVYLOTFDNELKRKTTELPQMGMDLKLVOLIAYT 236
Qy	304 LVKELAQVNEVHPYTVRKDALPEFTDYNQMDYALLNKSATGVFTDPDTGVEFLK 361	Qy	244 DWKETQKDPKGTYWNNYDMMFKPGAMAEVVKYADGVGPQWMLVNKEESKPDNIVTP 303
Db	296 MVQDAHQNKVWHPYTVRADQDLPDVTDNQLYDLYNKAGVDGLFTDPDKAVMFLQ 353	Db	237 DNNETQKQPKGRWVNNYDMMFKQAMKVAEYDGIGDYYHLY-AE8STKRNKLTG 295
RESULT 13			
Q8 2556	PRELIMINARY;	PRT;	356 AA.
AC	Q8Z556; Q7CB71;		
DT	01-MAR-2002 (TREMBLrel. 20, Created)	DT	07MAY7 PRELIMINARY;
DT	25-OCT-2004 (TREMBLrel. 28, Last annotation update)	AC	Q7MAY7 PRT;
DE	Glycerophosphoryl diester phosphodiesterase periplasmic (EC 3.1.4.46)	DT	01-MAR-2004 (TREMBLrel. 26, Last sequence update)
GN	Name=glpQ; OrderedLocusNames=STY2511, t0582;	DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonellae.	DE	Glycerophosphoryl diester phosphodiesterase, periplasmic
OC	Salmonella typhi; Enterobacteriaceae; Enterobacteria; Enterobacteriales; Enterobacteriaceae; Photorhabdus.	DB	(Glycerophosphoryl diester phosphodiesterase).
OX	NCBI_TaxID=601;	GN	Photorhabdus luminescens (subsp. laumontii).
RN	[1]	OS	Bacteria; Photobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Photorhabdus.
RP	SEQUENCE FROM N.A.	OC	NCBI_TaxID=141679;
RX	STRAIN=CT18;	OX	[1]
RA	MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;	RN	RN
RA	Parkhill J., Dougan G., James K.D., Thomson N.R., Pickett D., Wain J.,	RP	SEQLNDF FROM N.A.
RA	Baker C.M., Mungall K.L., Bentley S.D., Holden M.R., Sebastian N.,	RC	STRAIN=CT18;
RA	Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,	RX	MEDLINE=22957627; PubMed=14528314;
RA	Cronin A., Davis P., Davies R.M., Dowd L., White N., Parrat J.,	RA	TAOCHAUD E., RUBENIOL C., FRANGUEUL L., BUCHRIESE C., GIVAUDAN A.,
RA	Feltwell T., Hamlin T.S., Haque A., Hien T.T., Holoyd S., Jagels K.,	RA	TAOCHAUD E., RUBENIOL C., BOURSAUX-EUDE C., CHANDLER M., CHARLES J.-F.,
RA	Krogh A., Larsen T.S., Leather S., Monle S., O'Gaoar P., Parry C.,	RA	DASSA E., DEROSSE R., DESZELLE S., FREYSSINET G., GAUDRIAU S.,
RA	Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,	RA	MEIGNEAU C., LANOIS A., POWELL K., SINGER P., VINCENT R., WINGET V.,
RA	Whitehead S., Barrell B.G.;	RA	ZOUINE M., GLASER P., BEMARE N., DANCHIN A., KUNST F.,
RT	• "Complete genome sequence of a multiple drug resistant Salmonella	RT	"The genome sequence of the entomopathogenic bacterium Photorhabdus luminescens,"
RT	enterica serovar Typhi CT18.";	RT	Nature 413:848-852 (2001).
RN	[2]	RL	Nature Biotechnol. 21:1307-1313 (2003).
RP	SEQUENCE FROM N.A.	DR	EMBL; BX571872; CAB16432.1; -.
RC	STRAIN=CT18 / ATCC 700931;	DR	PHOTOLIB; plu4120; -.
RX	MEDLINE=22531367; PubMed=12644504;	DR	GO; GO:0008889; P: glycerophosphodiester phosphodiesterase act. . . ; IEA.
RX	DOI=10.1128/JB.185.7.2330-2337.2003;	DR	GO; GO:0006071; P: glycerol metabolism; IEA.
RA	Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,	DR	INTERPRO; IPR004129; GDD.
RA	Burland V., Kodoyiannaki V., Schwartz D.C., Blattner F.R.;	DR	Pfam; PF03009; GDDP. 1.
RT	"Comparative genomics of Salmonella enterica serovar Typhi strains Ty2 and CT18,"	KW	Complete proteome.
RL	J. Bacteriol. 185:2330-2337 (2003).	SQ	SEQUENCE 358 AA; 41182 MW; OBBB12D55785ABC5 CRC64;
DR	EMBL; AL622724; CAD07514.1; -.	Query Match	48.9%; Score 1210.5; DB 2; Length 358;
DR	EMBL; AB08836; AA08288.1; -.	Best Local Similarity	62.0%; Pred. No. 6.4e-76; Missmatches 54; Indels 11; Gaps 2;
DR	GO; GO:0008889; P: glycerophosphodiester phosphodiesterase act. . . ; IEA.	Matches 222;	Conservative 54;
DR	GO; GO:0016787; P: hydrolase activity; IEA.	4	KTLAISLLAAGVLAGCSSHSSNMANQMSDKIIIAHRCGAGYLDEHTLESKALAPAQAA 63
DR	GO; GO:0006071; P: glycerol metabolism; IEA.	6	KTMLIGIL-----TSMSMGIAQADKLVIAHRCGAGYLDEHTLPAKAMVYAEGA 55
DR	Pfam; PF03009; GDDP. 1.	64	DYLEQDLANTKDKGRIVVHDFLQGLTDVAKFPLKSHFRHTFEDEIEFQGLKSTGKVGTYPEIKAPWFFH
KW	Complete Proteome; Hydrolase.	56	DYLEQDLVMTKDKDEILVADHYLDRVTDVANKPFRARQDGRYIAIDFTLSEIKSLKPF 115
SQ	SEQUENCE 356 AA; 40437 MW;	Qy	124 NFETMDGKQAOQVPRFLPKSHPRHTFEDEIEFQGLKSTGKVGTYPEIKAPWFFH
DR	6 KNLNSVLMAGMTTGSAA-----VAAEKVVAHRCGAGYLDEHTLESKALAPAQAA 56	Db	116 GFDKNDRQIQNFSRFPWKSDPFRHTFOEIEVQGLNKSTGKVGTYPEIKAPWFFH 175
Qy	4 KTLAISLLAAGVLAGCSSHSSNMANQMSDKIIIAHRCGAGYLDEHTLESKALAPAQAA 63	Qy	184 QNGKDIAAETLKVKKYGYDKKCTDMVYLOTFDNELKRKTTELPQMGMDLKLVOLIAYT 243
Db	6 KNLNSVLMAGMTTGSAA-----VAAEKVVAHRCGAGYLDEHTLESKALAPAQAA 56	Db	176 KEGKDIISTKVAVLKAYG7TKSKDPLKGMDLKLVOLIAYT 235
Qy	64 DYLEQDLANTKDKGRIVVHDFLQGLTDVAKFPLKSHFRHTFEDEIEFQGLKSTGKVGTYPEIKAPWFFH 123	Db	244 DWKETQKDPKGTYWNNYDMMFKPGAMAEVVKYADGVGPQWMLVNKEESKPDNIVTP 303
Db	57 DYLEQDLVMTKDKDEILVADHYLDRVTDVANKPFRARQDGRYIAIDFTLDEIKSLKFT 116	Qy	126 DNNETYEKQSDGRWVNNYDMMFKQAMKVAEYDGIGDYYHLY-AE8STKRNKLTN 294
Qy	124 NFETMDGKQAOQVPRFLPKSHFRHTFEDEIEFQGLKSTGKVGTYPEIKAPWFFH 183	Db	117 GFDIENGKVKVOTYGRFPGRMGSDFRHTFEIEFQGLNHSRGNIGTYPEIKAPWFFH 176
Db	117 GFDIENGKVKVOTYGRFPGRMGSDFRHTFEIEFQGLNHSRGNIGTYPEIKAPWFFH 176	Qy	304 LVKELAQVNEVHPYTVRKDALPEFTDYNQMDYALLNKSATGVFTDPDTGVEFLK 361

Db 295 LVKEAHTNLEVHPYTFIRDQLPKYATSGDQLFDIYTNQAGVDFGVFTDPDGVKFLQ 352

RESULT 15

Q66FX7 PRELIMINARY; PRT; 371 AA.

ID Q66FX7

AC Q66FX7

DT 25-OCT-2004 (TREMBLrel. 28, Created)

DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)

DE Glycerophosphoryl diester phosphodiesterase precursor (EC 3.1.4.-6).

GN Name=gIPQ; ORFNames=YPTB0208;

OS Yersinia pseudotuberculosis IP 32953.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Yersinia.

OX NCBI_TaxID=273123;

RN [1]

RP SEQUENCE FROM N A.

RC STRAIN/PI 32953;

RX PubMed:15358858;

RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,

RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin L.V.,

RA Rubakker R.R., Fowler J., Hinnbusch B.J., Marceau M., Medigue C.,

RA Simonet M., Chenal-Francisque V., Souza B., Dacheux D., Elliott J.M.,

RA Derbise A., Hauser L.J., Garcia B.;

RT "Insights into the genome evolution of *Yersinia pestis* through whole

RT genome comparison with *Yersinia pseudotuberculosis*,"

RL Proc. Natl. Acad. Sci. U. S. A. 101:13826-13831 (2004).

DR EMBL; BX93698; CAH19448.1;

DR GO: GO:0016787; F:hydrolyase activity; IEA.

DR InterPro; IPR004129; GDFD.

KW Hydrolase; Signal.

PT SIGNAL 1 25 MW; Potential 1.

SEQ ♦ SEQUENCE 371 AA; 42395 MW; 03F96587AAF6262A CRC64;

Query Match 48.8%; Score 1210; DB 2; Length 371;

Best Local Similarity 62.7%; Pred. No. 7.3e-76; Mismatches 75; Indels 6; Gaps 4;

Matches 227; Conservative 54; Mismatches 75;

Qy 4 KTLALSLLAAGVLAGCSSHSSNANTOMKS ---DKLIAHKGASGVLPEHTLESKALAFKA 60

Db 5 KTLALSLLAAGVLAGCSSHSSNANTOMKS ---DKLIAHKGASGVLPEHTLESKALAFKA 60

Qy 121 MTENFE-TKDQKQAAQVYNRPPPLWKSHEFRIITFEDELFQGLESTGKVGIVYPEIKAP 179

Db 125 FTEGFDDKNGKCKVQSYPNRFPMGKSDPRTVHTFOQSELEFIQGLNISTGKNTGYYPEIKAP 184

Qy 180 WPHHQNGKQDIAAETLKVKKYDQKTDMYLQTFDNEELRIKTELLPONGMDLKLVQL 239

Db 185 WPHHQBSKDKISTKVKLETLKEYTNKDKVTLQSDPVNELRKIQNELEPKNGMDLKLVQL 244

Qy 240 IAYTDWKETQBDPKGKCYWVNTYNDMMKPGAMAEVVYKAQDGVPGKMYMLVNEESEKPDNI 299

Db 245 IAYTDWNETYQQLDGRKWNVNTYNDMMKPGAMKEYAHADGIGPDYHMLVVKTSKAD-1 303

Qy 300 VYPLVKELAQVNENVHPYTFIRDQLPKYATDINQLFDIYTNQAGVDFGVFTDPDGVKFLQ 359

Db 304 KLTDLVKEAHASNNMVFPTIRDRPKYATDINQLFDIYTNQAGVDFGVFTDPDGVKFLQ 363

Qy 360 LK 361

Db 364 LQ 365

THIS PAGE BLANK (USPTO)